Journal of
Marketing

January 2008
Volume 72, Number 1

CUK- H05817-16-KP2122

Preventing Digital Music Piracy: The Carrot or the Stick? 1 Rajiv K. Sinha & Naomi Mandel

Marketing Initiatives, Expected Cash Flows, and Shareholders' Wealth 16 Ramesh K.S. Rao & Neeraj Bharadwaj

Interaction Orientation and Firm Performance 27 Girish Ramani & V. Kumar

Expanding Business-to-Business Customer Relationships: Modeling the Customer's Upgrade Decision 46 Ruth N. Bolton, Katherine N. Lemon, & Peter C. Verhoef

Chief Marketing Officers: A Study of Their Presence in Firms' Top Management Teams 65 Pravin Nath & Vijay Mahajan

Understanding Service Retention Within and Across Cohorts Using Limited Information 82 David A. Schweidel, Peter S. Fader, & Eric T. Bradlow

When Old Is Gold: The Role of Business Longevity in Risky Situations 95 Preyas S. Desai, Ajay Kalra, & B.P.S. Murthi

The Readability of Marketing Journals: Are Award-Winning Articles Better Written? 108 Alan G. Sawyer, Juliano Laran, & Jun Xu

Stage-Gate Controls, Learning Failure, and Adverse Effect on Novel New Products 118 Rajesh Sethi & Zafar Iqbal

16

AMERICAN
MARKETING
ASSOCIATION

Call for Papers

Journal of Public Policy & Marketing **Spring 2009**

Consumption Constraints:
Implications for Consumer Welfare and Public Policy

Special Issue Editor:
Stacey Menzel Baker
University of Wyoming

The vast majority of the world does not live like the people whose lives are depicted on television, in movies, in glossy magazines, or on the Internet. Nor is the vast majority of the world depicted in research samples in the academic literature. Four-fifths of people in the world live in poverty, one of five people in the United States has some kind of disability, and one of five people in the United States is functionally illiterate. Half of the people in the world have never used a telephone, and fewer than one-fifth have access to the Internet. In other words, many people are constrained in their consumption. These constraints can be experienced differently by different people, and sometimes these constraints can contribute to consumer vulnerability.

People may be constrained in their consumption by any number of things, including, but not limited to, the following:

- Skills and abilities (e.g., physical disabilities),
- Cognitive capacity (e.g., mental disabilities),
- Financial resources,
- Environmental factors (e.g., logistical barriers in retail and service settings),
- Dominant worldviews,
- Lack of social networks,
- Geographic or social isolation,
- Regulatory or economic policies,
- Poor infrastructure,
- Fragmented markets,
- Political instability (e.g., war),
- Access to choices (e.g., assortments),
- Access to goods and services, and
- Access to information.

The nature of these constraints may vary for any number of reasons, including, but not limited to, the following:

- Temporary versus long run,
- Public versus private,
- Justifiable restrictions versus nonjustifiable restrictions,
- Self-induced versus externally induced, and
- Necessity versus luxury.

KP 2722

Journal of Public Policy & Marketing (*JPP&M*) invites scholarly articles that address consumption constraints. Articles may be theoretical or empirical and may employ traditional or innovative research methods that tackle compelling issues that are relevant to consumption constraints. Research for this issue of *JPP&M* should have the potential to reduce consumers' vulnerabilities, transform consumers' lives, and contribute to public policy.

Papers will be reviewed in accordance with *JPP&M* guidelines; for further details, including manuscript guidelines and submission instructions, see www.marketingpower.com/jppm. Manuscript length is typically 30–40 double-spaced pages, including references, tables, and figures (12-point font); the maximum length is 50 pages. For further information, contact the special issue editor, Stacey Menzel Baker, at smbaker@uwyo.edu. The manuscript submission deadline is **April 1, 2008.**

Journal of
Marketing

Vol. 72, No. 1

AMERICAN MARKETING ASSOCIATION

University
16.4.08
ALIPORE
CAMPUS

EDITOR
Roland T. Rust
Robert H. Smith School of Business
3451 Van Munching Hall
University of Maryland
College Park, MD 20742-1815

ASSISTANTS TO THE EDITOR
Chris Bartone
E-mail: jom@ama.org

Millicent Locke
University of Maryland

WEB EDITOR
Siva K. Balasubramanian
Southern Illinois University

BOOK REVIEWS
Ramarao Desiraju
University of Central Florida

VICE PRESIDENT—PUBLICATIONS
Russell S. Winer
New York University

CHIEF EXECUTIVE OFFICER
Dennis Dunlap

GROUP PUBLISHER
Jack Hollfelder

EDITORIAL STAFF
Phone: (312) 542-9000
Fax: (312) 542-9001

MANAGING EDITOR
Francesca Van Gorp Cooley

TECHNICAL EDITOR
Andy Seagram

GRAPHICS EDITOR
Mary Loye

Advertising Staff
Fax: (312) 922-3763

Richard Ballschmiede, Advertising
(312) 542-9076

Sally Schmitz, Production
(312) 542-9038

JM Web site:
Submissions: www.marketingjournals.org/JM
Main page: www.marketingpower.com/JM

The *Journal of Marketing* (ISSN 0022-2429) is published bimonthly (January, March, May, July, September, and November) *Annual Subscription Rates*: Subscriptions are $115 for individuals in the United States and Canada and $160 in other countries. Library, school, and business subscriptions are $325 in the United States and Canada and $360 in other countries. Subscriptions for AMA members are $60. Prices are for one year. Single copy rate: $50 individual, $60 corporate in the United States and Canada; add $5 in other countries. Prepayment is required in U S. dollars or the equivalent.

Published by the Publications Group of the American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266. Periodicals postage paid at Chicago, IL and additional mailing offices. POSTMASTER: Send address changes to the *Journal of Marketing*, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266, USA.

•MANUSCRIPTS: Electronic submissions must be made via the *JM* Web site at http://www.marketingjournals.org. Paper submissions are not accepted Editorial policies, review board members, Tables of Contents for *JM*, and so on, are available on the *JM* Web site at http //www.marketingpower.com/jm.

•ADVERTISERS AND ADVERTISING AGENCIES assume liability for all content (including text, representations, and illustrations) of advertisements published, and also assume responsibility for any claims arising therefrom made against the publisher The right is reserved to reject any advertisement.

•SUBSCRIPTIONS*:* Members: Send to American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266 (1-800-AMA-1150 or 312/542-9000), or e-mail to info@ama.org. Nonmembers: Send to American Marketing Association, P.O. Box 11806, Birmingham, AL 35202, 1-800-633-4931, 205-995-1567, Fax: 205-995-1588, e-mail: amasubs@ebsco.com. Canadian residents add 6% GST (GST registration # 12747852).

•PERMISSIONS AND REPRINT INFORMATION: Without written permission from the AMA, any copying or reprinting (except by authors reprinting their own work) is prohibited. Requests for permission to reprint—such as copying for general distribution, advertising, or promotional purpose; for creating new collective works; or for resale—should be submitted in writing to permissions@ama.org. Bulk reprints are available by contacting Tamara Groft, Sheridan Reprints, (800) 352-2210. Permission is granted by AMA for those registered with the Copyright Clearance Center (CCC), 222 Rosewood Dr., Danvers, MA 01923, to photocopy any article herein for a flat fee. Contact CCC with ISSN 0022-2429, volume, and first and last page numbers of each article copied.

The views expressed in *JM* are not necessarily those of the editorial staff or the publisher. **All articles in this journal, unless otherwise noted, have undergone a blind review process.**

Copyright © 2008, American Marketing Association. All rights reserved

Editorial Review Board

ROHINI AHLUWALIA
University of Minnesota

KUSUM L AILAWADI
Dartmouth College

JOSEPH ALBA
University of Florida

TIM AMBLER
London Business School

EUGENE W ANDERSON
University of Michigan

TOR WALLIN ANDREASSEN
Norwegian School of Management

ASIM ANSARI
Columbia University

ERIC J. ARNOULD
University of Wyoming

KWAKU ATUAHENE-GIMA
CEIBS

BARRY L BAYUS
University of North Carolina

WILLIAM BEARDEN
University of South Carolina

PAUL D BERGER
Boston University

PAUL N. BLOOM
Duke University

LISA E BOLTON
University of Pennsylvania

RUTH N. BOLTON
Arizona State University

ROBERT BORDLEY
General Motors Corporation

DOUGLAS BOWMAN
Emory University

RODERICK J. BRODIE
University of Auckland

STEVEN P. BROWN
University of Houston

JOSEPH CANNON
Colorado State University

DIPANKAR CHAKRAVARTI
University of Colorado

PIERRE CHANDON
INSEAD

YUXIN CHEN
New York University

ALEXANDER CHERNEV
Northwestern University

PRADEEP K CHINTAGUNTA
University of Chicago

KEVIN CLANCY
Copernicus Marketing Consulting

PETER J DANAHER
University of Melbourne

KO DE RUYTER
Maastricht University

JOHN A DEIGHTON
Harvard University

MARNIK G DEKIMPE
Tilburg University and Catholic
University Leuven

WAYNE S DESARBO
Pennsylvania State University

RAVI DHAR
Yale University

NAVEEN DONTHU
Georgia State University

TULIN ERDEM
New York University

PETER S FADER
University of Pennsylvania

GARY L. FRAZIER
University of Southern California

GARY F. GEBHARDT
University of South Florida

INGE GEYSKENS
Tilburg University

PETER N GOLDER
New York University

RAJDEEP GREWAL
Pennsylvania State University

ABBIE GRIFFIN
University of Utah

SUNIL GUPTA
Harvard University

ZEYNEP GURHAN-CANLI
KOC University

MICHAEL HAENLEIN
ESCP-EAP

JOHN R HAUSER
Massachusetts Institute of
Technology

JAN B HEIDE
University of Wisconsin–Madison

CHRISTIAN HOMBURG
University of Mannheim

WAYNE D. HOYER
University of Texas at Austin

JOEL HUBER
Duke University

JOHN HULLAND
University of Pittsburgh

G TOMAS M. HULT
Michigan State University

J. JEFFREY INMAN
University of Pittsburgh

GANESH IYER
University of California, Berkeley

SANJAY JAIN
Texas A&M University

CHRIS JANISZEWSKI
University of Florida

SANDY D JAP
Emory University

BERNARD J JAWORSKI
Monitor Group

P.K. KANNAN
University of Maryland

TIMOTHY L. KEININGHAM
Ipsos Loyalty

AMNA KIRMANI
University of Maryland

AJAY K KOHLI
Emory University

ROBERT V KOZINETS
York University

V. KUMAR
University of Connecticut

DONALD R LEHMANN
Columbia University

KATHERINE N. LEMON
Boston College

MARY FRANCES LUCE
Duke University

XUEMING LUO
University of Texas at Arlington

ROBERT F LUSCH
University of Arizona

JOHN G LYNCH JR.
Duke University

VIJAY MAHAJAN
University of Texas at Austin

RICHARD G MCFARLAND
Kansas State University

CARL F. MELA
Duke University

VIKAS MITTAL
University of Pittsburgh

WENDY MOE
University of Maryland

CHRISTINE MOORMAN
Duke University

JOHN MOWEN
Oklahoma State University

EITAN MULLER
New York University and
Tel-Aviv University

SCOTT A. NESLIN
Dartmouth College

RICHARD NETEMEYER
University of Virginia

RAKESH NIRAJ
University of Southern California

STEPHEN M. NOWLIS
Arizona State University

JOSEPH C. NUNES
University of Southern California

A PARASURAMAN
University of Miami

KOEN PAUWELS
Dartmouth College

WILLIAM D. PERREAULT
University of North Carolina

ROBERT PETERSON
University of Texas at Austin

F.G.M. PIETERS
Tilburg University

JAIDEEP PRABHU
Imperial College

WILLIAM P PUTSIS
University of North Carolina

ARVIND RANGASWAMY
Pennsylvania State University

AKSHAY RAO
University of Minnesota

VITHALA RAO
Cornell University

BRIAN T. RATCHFORD
University of Texas at Dallas

WERNER REINARTZ
INSEAD

ARIC RINDFLEISCH
University of Wisconsin–Madison

FRED SELNES
Norwegian School of Management

STEVEN M. SHUGAN
University of Florida

ITAMAR SIMONSON
Stanford University

STANLEY F. SLATER
Colorado State University

REBECCA SLOTEGRAAF
Indiana University

JOYDEEP SRIVASTAVA
University of Maryland

JAN-BENEDICT E M. STEENKAMP
University of North Carolina

DAVID W. STEWART
University of Southern California

STEFAN STREMERSCH
Erasmus University Rotterdam

BAOHUNG SUN
Carnegie Mellon University

GERARD J TELLIS
University of Southern California

JACQUELYN THOMAS
Northwestern University

CRAIG THOMPSON
University of Wisconsin–Madison

DEBORA THOMPSON
Georgetown University

CHRISTOPHE VAN DEN BULTE
University of Pennsylvania

STIJN VAN OSSELAER
Erasmus University Rotterdam

P RAJAN VARADARAJAN
Texas A&M University

STEPHEN L VARGO
University of Hawaii, Manoa

SAJEEV VARKI
University of South Florida

RAJKUMAR VENKATESAN
University of Virginia

PETER VERHOEF
University of Groningen

KENNETH WATHNE
University of Wisconsin–Madison

MICHEL WEDEL
University of Maryland

BARTON A WEITZ
University of Florida

RUSSELL S. WINER
New York University

CAROLYN YOON
University of Michigan

ANTHONY ZAHORIK
Burke Institute

VALARIE A ZEITHAML
University of North Carolina,
Chapel Hill

Z JOHN ZHANG
University of Pennsylvania

JIE ZHANG
University of Maryland

KP2722

Rajiv K. Sinha & Naomi Mandel

Preventing Digital Music Piracy: The Carrot or the Stick?

The goal of this article is to ascertain the factors that govern consumers' willingness to pirate a digital product, such as a digital music track. The authors assess the tendency to pirate with both indirect measures (e.g., willingness to pay for the legal alternative) and direct measures (e.g., piracy preference). Whether measured indirectly or directly, the tendency to pirate depends, to different extents, on three key factors: positive incentives (e.g., improved functionality of the legal Web site), negative incentives (e.g., perceived risk of piracy), and consumer characteristics. Based on three studies, the results suggest that negative incentives are a strong deterrent for certain consumers but can actually increase piracy tendencies for others. Conversely, positive incentives, such as improved functionality, can significantly reduce the tendency to pirate among all the consumer segments studied. The authors conclude by discussing prescriptive recommendations for the recording industry.

Keywords: music, piracy, willingness to pay, risk tolerance

In recent years, music industry sales have declined in the United States by an average of 7% per year, after peaking in 1999 (Connolly and Krueger 2005). Much of this sales decline has been attributed to the widespread use of illegal file-sharing sites, such as Napster (introduced in 1999 and shut down in 2000) and, more recently, Grokster, LimeWire, and Morpheus (Liebowitz 2004; Peitz and Waelbroeck 2004). However, the success of iTunes and other legal file-sharing Web sites (Wade 2004) seems to indicate that though many consumers pirate their music, a portion of music consumers are willing to pay a positive amount to download music legally.

The primary goal of this article is to ascertain the factors that govern a consumer's decision to pirate a digital music track. Because of the sensitive nature of survey questions about illegal activity, we measure participants' willingness to pirate both directly (through a piracy preference measure) and indirectly (through willingness to pay [WTP] for the legal alternative). Whether it is measured indirectly or directly, we find that the tendency to pirate depends, to different extents, on three key factors: positive incentives (e.g., improved functionality of the legal Web site), negative incentives (e.g., perceived risk of piracy), and consumer characteristics.

Our findings are highly relevant to marketers of music and other products that might be digitized, such as movies and software. Record companies have used much of their resources in the past eight years to fight digital piracy in court. Their numerous copyright infringement lawsuits are intended to produce fear and restraint in consumers; however, there is little evidence that these lawsuits have directly decreased piracy levels. Although illegal downloading has decreased somewhat in recent years (Wade 2004), it is unclear whether this decrease is due to the lawsuits brought about by the Recording Industry Association of America (RIAA) or to the recent availability of rich, user-friendly, legal Web sites, such as iTunes and Real Rhapsody, as an alternative to illegal downloading (Easley, Michel, and Devaraj 2003). Although the music industry is currently working with file-sharing companies, such as Grokster, to convert free trading of digital music to paid services (Hansell 2005), there is much disagreement among industry players about issues such as the price to charge (Bhatia, Gay, and Honey 2003) and whether features such as digital rights management actually deter digital piracy (as intended) or encourage it (BBC News 2007; Bechtold 2004).

In this article, we develop hypotheses, drawn from existing economic, psychology, and consumer theory, that predict the relative success of negative incentives (e.g., legal enforcement) and positive incentives (e.g., enhanced Web features) in decreasing piracy and increasing WTP for the legal product. Our results from three studies suggest that negative incentives are a strong deterrent for certain consumers but can actually increase piracy tendencies for others. Conversely, positive incentives, such as improved functionality, can significantly reduce the tendency to pirate among all the consumer segments we studied.

Conceptual Framework

Digital music is frequently characterized as an information good that is expensive to produce but cheap to reproduce. In perfectly competitive markets, prices of information goods have been predicted to fall to zero (Shapiro and Varian 1999, pp. 19–23). Consistent with this notion, the decreased search costs in an online environment have been found to

---

Rajiv K. Sinha is Associate Professor of Marketing (e-mail: rajiv.sinha@asu.edu), and Naomi Mandel is Associate Professor of Marketing (e-mail: naomi.mandel@asu.edu), W.P. Carey School of Business, Arizona State University.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

increase consumers' price sensitivity, but only for commonly available merchandise, not for unique or customizable products (Lynch and Ariely 2000). Unfortunately for record companies, a popular song is commonly viewed as a commodity that can be easily copied with little loss in sound quality, making it a strong candidate for price erosion. We propose that a consumer's decision of whether to download a digital song for free (i.e., to pirate) or to pay for it (and, by extension, how much to pay for it) depends on three key factors: negative incentives (e.g., perceived risk of piracy), positive incentives (e.g., improved functionality of the legal Web site), and consumer characteristics.

Negative Incentives: Perceived Risk of Piracy

A consumer's choice to participate in a risky behavior, such as illegal file sharing, depends on the perceived risk associated with that behavior, which can be broken down into two dimensions: uncertainty (e.g., the probability of getting caught) and adverse consequences (e.g., the punishment) (Bauer 1960). In general, people are risk averse, and they are more averse to losses than to equivalent gains (Kahneman and Tversky 1984). Therefore, it seems plausible that increasing either the probability of getting caught or the negative consequences of illegal downloading should lead to a decrease in piracy. The RIAA has attempted to increase the perceived probability of getting caught by suing more than 14,000 file swappers (Reuters 2005). It has also attempted to increase the perceived magnitude of negative consequences by imposing excessive penalties, such as a fine of $750 per copyright infringement, which would amount to approximately $3 million for someone who has accumulated 4000 songs on a hard drive (Barker 2004). Indeed, some of the RIAA's actions appear to be working. For example, the NPD Group found that approximately 1.4 million people have deleted all illegal MP3 files from their computers in an attempt to protect themselves from impending lawsuits (Wade 2004). Therefore, we expect that there is a significant effect of perceived risk on consumers' tendency to pirate.

> H_1: A consumer's perceived risk of getting caught downloading illegally decreases his or her tendency to pirate a digital song.

A closely related, albeit distinct, construct (to piracy potential) is the WTP for a legal song. The distinction between the two is that all pirates are unwilling to pay for a product, but not all consumers with a zero WTP (or a WTP that is less than the legal retail price) can be considered pirates. Rather, some of these consumers may prefer to purchase the music in CD format or to forgo the purchase because their WTP is less than the retail price. Nevertheless, the average WTP of any consumer segment will decrease with an increase in piracy so that ascertaining the average WTP of a group of consumers who are interested in owning a particular digital track (legally or illegally) provides an important indication of their tendency to pirate. For example, in the context of H_1, because fewer people should be willing to pay $0 (or pirate) when they believe that they are likely to be caught, the average WTP among participants should increase. Therefore, we expect that WTP for a legal download increases when perceived risk increases.

Consumer Characteristics

Next, we examine the demographic and psychographic profile of a consumer who downloads music illegally. Moe and Fader (2001) suggest that for "hedonic portfolio" (i.e., collectible) products, such as music, several different consumer segments exist, and these segments behave differently in terms of their product acquisition behavior. How do digital pirates differ from consumers who either download legally or purchase music in the traditional CD format? Digital music pirates tend to be young and male, and their likelihood of pirating increases with the price of the song, the popularity of the song, and the size of the bandwidth available (Bhattacharjee, Gopal, and Sanders 2003). College students make up a large portion of illegal downloaders, as a result of high-speed network connections and an abundance of free time (Wade 2004). Consistent with this argument, CD sales declined more at stores near universities after Napster's appearance than did sales in other regions (Fine 2000).

We propose that a consumer's optimal stimulation level (OSL) (Leuba 1955; Raju 1980) also determines his or her piracy potential. According to OSL theory, every person has an ideal level of stimulation; when the stimulation offered by the external environment is too low, he or she will attempt to increase stimulation, and when the external stimulation is too high, he or she will attempt to decrease it. A consumer's OSL is typically measured with scales such as the sensation-seeking scale (Zuckerman 1979) or the arousal-seeking-tendency scale (Mehrabian and Russell 1974; Raju 1980). People with high OSLs desire novel, varied, and complex experiences and sensations, and in general, they are willing to take the risks necessary to achieve such experiences (Zuckerman 1979). The OSL is highly correlated with demographic variables, such as age, gender, employment status, and education. High-OSL consumers are relatively younger, more educated, better employed, and more likely to be male than low-OSL consumers (Raju 1980; Zuckerman 1994, p. 114). Finally, the OSL is an important determinant of a consumer's exploratory behaviors (Raju 1980). High-OSL consumers are more likely to participate in behaviors such as innovativeness, risk taking, and variety seeking, as measured by Raju's (1980) exploratory tendencies in consumer behavior (EXPL) scale.

High-OSL consumers show a greater preference for complexity, asymmetry, and the appearance of movement in visual images (Zuckerman 1994, p. 201). They also prefer complexity and novelty in music and enjoy live music performances (Schierman and Rowland 1985). Because high-OSL consumers are likely to seek out new artists and songs that are not currently in their collections and are willing to take risks to do so, we believe that OSL is positively related to the propensity to download illegally.

> H_2: A consumer's OSL is positively related to his or her tendency to pirate a digital song.

Along the same lines, we expect that WTP for a legal download decreases when OSL increases.

Furthermore, we expect that OSL moderates the relationship between perceived risk and piracy. Recent studies indicate that legal restrictions may not always inhibit illegal file swapping; on the contrary, when consumers believe that the rules are too excessive, they become even more likely to break them (Raghu et al. 2005). Therefore, it is unclear whether increasing the perceived risk of piracy either increases or decreases consumers' digital piracy. We propose that the effect of perceived risk depends on a consumer's OSL. Specifically, low-OSL consumers should be less likely to pirate when the perceived risk is increased. However, high-OSL consumers have a higher tolerance for risk; therefore, we expect that increasing the perceived risk will not decrease their piracy levels. Indeed, if taking such a risk provides a higher level of arousal for high-OSL consumers, it might actually increase their piracy levels.

> H_3: OSL moderates the relationship between perceived risk and piracy. For low-OSL consumers, increasing the risk of being caught decreases piracy, and for high-OSL consumers, increasing the risk of being caught does not decrease piracy.

Similarly, we expect that increasing perceived risk increases WTP for low-OSL consumers but not for high-OSL consumers.

We also examine the role of ethics in the file-sharing decision. Previous research has found that people who download illegally have a lower level of ethical concern than those who do not (Levin, Dato-on, and Rhee 2004). Thus, we expect that consumers who score high on the Machiavellian scale (Christie and Geis 1970) are less willing to pay for legal downloads than those who score low. Machiavellianism is characterized by "cool detachment" and clever manipulation of others to achieve personal goals. Therefore, it is likely that high scorers on this scale demonstrate less sympathy for the artists' and music companies' losses as a result of music piracy. Furthermore, respondents who score high on Machiavellianism demonstrate a distrust of others. For example, the Machiavellian scale (Mach IV; Christie and Geis 1970) includes items such as "It is safest to assume that all people have a vicious streak and it will come out when they are given a chance" and "Anyone who completely trusts anyone else is asking for trouble." We reasoned that many consumers engage in cheating behaviors, such as illegal downloading, because of the belief that "everyone else is doing it." Furthermore, it seems intuitive that high scorers on the Machiavellian scale will demonstrate a higher OSL than low scorers. Therefore, we expect that high scorers on the Machiavellian scale are willing to pay less than low scorers for a legal download.

> H_4: Consumers with high scores on the Machiavellian scale demonstrate a greater tendency toward digital piracy than those with low scores.

Positive Incentives: Improved Functionality

Recent decreases in piracy may not be due to the RIAA's legal actions but rather to the availability of attractive legal alternatives, such as iTunes, which was launched in 2003 and accounted for more than 25 million legal downloads by the end of that year (Wade 2004). A way to mitigate the downward pressure on prices for pure digital products is to differentiate the product or service and create different versions for diverse market segments (Shapiro and Varian 1999, p. 53). Bakos and Brynjolfsson (1999, 2000) recommend using bundling strategies for information goods with zero marginal cost. For example, online music retailers might bundle songs by offering a monthly subscription for unlimited downloads (e.g., Yahoo Music Unlimited) or customized mix CDs (e.g., Real Rhapsody). Alternatively, they might charge higher prices for more popular songs, live recordings, or out-of-print recordings. For example, Sony BMG recently introduced ShopBootlegs.com, where consumers can legally download live concert recordings for $6.98 to $13.98 (Smith 2005).

In addition, online music retailers have been encouraged to differentiate from competitors and increase the perceived value of their products by offering superior service, user-friendly features, or customized recommendations (Bhatia, Gay, and Honey 2003; Easley, Michel, and Devaraj 2003). In support of this idea, a Gartner survey of 1500 "early adopters" of the Internet indicated that one in three consumers would be interested in using recommendation engines to suggest new music to try, based on their existing preferences (Sullivan 2005). Furthermore, navigation design that is easy to learn leads to more frequent visits and higher levels of purchasing (Johnson, Bellman, and Lohse 2003). Therefore, we propose that Web sites that offer features such as recommendation engines or easy navigation are able to command a higher price per song (thus resulting in lower levels of piracy) than Web sites that do not offer such features.

> H_5: Enhanced functionality on a legal Web site decreases a consumer's tendency to pirate a digital song.

Piracy and Social Norms

The threat of embarrassment or shame associated with being caught might also influence a consumer's tendency to pirate. Reports of people arrested or sued by the RIAA have appeared in local and national newspapers, in television broadcasts, and on numerous Web sites and blogs. Breaking the law represents a transgression of social norms, which has been shown to cause embarrassment (Keltner and Buswell 1997). The negative outcome resulting from getting caught might threaten a consumer's identity, face, or approval from others (Weinstein and Martin 1969). Therefore, we expect that when the expected level of embarrassment associated with getting caught increases, people are more willing to pay for a legal download.

Another variable related to the level of embarrassment associated with illegal activity is the question whether the downloaded music is consumed publicly or privately. Embarrassment is a public emotion, and it usually occurs when a consumer feels conspicuous or self-conscious in front of others (Fenigstein 1979; Miller 1992). Consumers may feel embarrassed when purchasing products intended to be used in private, such as condoms (Dahl, Manchanda, and Argo 2001). Consuming an embarrassing product carries a social risk, which in some cases can have a more powerful effect on behavior than a financial risk (e.g., Mandel 2003). Therefore, illegal downloading behavior might be viewed as embarrassing if the downloaded music is consumed with friends (e.g., at a party). Although the purchase price of a digital song may not be obvious to the party's guest, a pirated song is more likely to demonstrate skips, imperfections, abrupt endings, and/or lower digital quality. Therefore, consumers who are planning to play the music publicly might be more concerned about song quality and might be more willing to pay for a legal song, which is more likely to have been recorded at a higher bit rate. This line of reasoning leads to the following hypotheses:

H_6: The level of embarrassment associated with getting caught downloading illegally decreases a consumer's tendency to a pirate a digital song.

H_7: Consumers are more likely to pirate a song to be consumed privately than to pirate a song to be consumed publicly.

Because digital piracy has negative legal and social implications, it might be considered a consumer "dark-side variable" and thus may be subject to socially desirable responding (Mick 1996). Measures of dark-side variables, such as materialism, can be distorted in the direction of respondents' perceptions of "correct" or socially desirable responses. Therefore, it is important to triangulate our measures of piracy preferences to ensure that they are not biased by socially desirable responding. A possible technique for debiasing stated piracy preferences is to frame them in a more positive manner by estimating respondents' WTP for a digital song of their choice under different circumstances. A WTP that is significantly lower than the prevailing market price would imply a higher preference for piracy, particularly for downloading a desirable digital song, which is a low-budget item and not subject to the usual budget constraints. Because piracy preference is a measure that may be highly sensitive to social norms, in Study 1, we use WTP as a surrogate for the tendency to pirate.

However, as we pointed out previously, WTP is an imperfect measure of willingness to pirate because there are consumers who exhibit low WTP but are also unlikely to pirate. For example, some people prefer to purchase CDs rather than to download a digital track (whether legal or illegal), and others are not in the habit of purchasing music. Therefore, in Studies 2 and 3, we examine alternative measures of piracy. Another option for debiasing piracy measures is the use of indirect questioning (Fisher 1993). This technique requires participants to project their own feelings or behavior onto others by answering questions from the perspective of another person or group. Indirect questioning can reduce the influence of socially desirable responses on measures that are sensitive to social influence (Fisher 1993). Therefore, Studies 2 and 3 use several different piracy measures. Specifically, Study 2 uses a direct piracy measure and the WTP measure used in Study 1. Study 3 measures piracy preferences from three perspectives: (1) approval of a fellow student's piracy behavior, (2) likelihood of piracy by other students at the same university, and (3) likelihood of one's own piracy.

In the remainder of the article, we first develop a model to measure piracy indirectly through WTP. Next, we present three studies that use both indirect and direct measures of piracy to test our hypotheses.

Study 1 examines the effects of perceived risk (H_1), Machiavellianism (H_4), embarrassment (H_6), and public versus private consumption (H_7) on WTP. Study 2 explores the effects of perceived risk (H_1), OSL (H_2 and H_3), and enhanced functionality (H_7) on both WTP and the tendency to pirate. Finally, Study 3 further investigates the effects of perceived risk, OSL, and enhanced functionality on a consumer's own tendency to pirate, as well as approval and expectations of others' piracy behaviors.

A Model for Estimating WTP

Various techniques, such as conjoint analysis and logit models, have been proposed to measure WTP for products (Jedidi and Zhang 2002). However, in marked contrast to these techniques, which often use attribute-based methods and are designed for assessing the magnitude of payment, we were interested in piracy and the consumers' tendency to avoid paying for the product. More important, we wanted to investigate how we might convert nonpaying consumers to paying ones. In this sense, our problem is similar to that of assessing consumers' WTP for public goods. Pure digital or information goods (e.g., music) share the two defining characteristics of public goods; they are both nonrivalrous and nonexcludable. Both properties are a consequence of the notion that information can be freely copied. Copying music makes its consumption nonrivalrous because consumption by one consumer does not reduce the ability of others to consume the same product. Copying also makes music nonexcludable because it is not possible to prevent others from consuming the same good by paying for it. A popular method for measuring and analyzing the value of public goods is the contingent valuation method (Haab and McConnell 2002; Wertenbroch and Skiera 2002). As we outline subsequently, these models are ideally suited for assessing the piracy potential of information goods.

These models directly estimate the WTP function by using a choice format that asks respondents whether they would be willing to pay a stated price for the product in question. In one version of the contingent valuation method, called the double-bound dichotomous choice (DBDC) format (Mitchell and Carson 1989),[1] respondents are provided with a sequence of two bids and are asked whether their WTP for a desirable product of their choice equals or exceeds that bid. The magnitude of the second bid depends on the answer (yes/no) to the first bid. The initial bid is denoted as B_1, and respondents are asked whether they would purchase the product if it were priced at B_1. If the answer is yes, they are presented with a new bid, B_H, where

---

[1] Based on the seminal work of Bishop and Heberlein (1979), this method was endorsed by the National Oceanic & Atmospheric Administration panel in 1993 (Arrow et al. 1993). It provides incentives for the truthful revelation of consumer preferences and greatly simplifies the cognitive tasks faced by respondents, particularly compared with the stated preference choice modeling approaches for measuring WTP. The numerous recommendations made by the panel to enhance the validity of the DBDC procedure were followed in this study.

$B_H > B_1$, and they are asked whether they are willing to pay the higher amount B_H. However, if the respondent's initial response is negative, they are presented with $B_L < B_1$ and are asked whether they are willing to pay the lower amount B_L. By varying the bid across subsamples systematically, it is possible to calculate the proportion of respondents who are willing to pay as a function of each bid. As we outline in the Appendix, the DBDC model provides a simple method to estimate the WTP for a digital song, which is the first of our three alternative measures for assessing piracy.

The main advantage of this approach over the stated preference choice models and conjoint models is that it more closely resembles consumer choice in an actual market situation, in which consumers are given prices (as opposed to being faced with open-ended questions, attribute descriptions, or lotteries). In addition, because the stated prices vary across subsamples, no single respondent can influence the final outcome. As a result, these dichotomous choice techniques are less susceptible to gaming and strategic behavior (Mitchell and Carson 1989), as well as starting point bias (Smith, Olsen, and Harris 1999). Finally, given the explicit focus on whether respondents would "pay anything at all," this method is well suited for developing measures of willingness to pirate and for assessing the impact of covariates on piracy.

Study 1: The Effects of Machiavellianism, Risk, Embarrassment, and Public Versus Private Consumption on WTP

Pretest

We collected data from 425 students during fall 2002 to get initial WTP estimates. These respondents represented a mix of undergraduate students, graduate students, and executives. We collected the data in both a classroom setting and a popular campus music store. These students were randomly exposed to one of five bid sets, which varied systematically between $0 and $1.50; the following numbers denote the initial bid, the subsequent higher bid (if respondents said yes to the initial bid), and the subsequent lower bid (if respondents said no to the initial bid), respectively: ($.10, $.25, $.05), ($.25, $.50, $.15), ($.50, $.75, $.25), ($.75, $1.00, $.50), and ($1.00, $1.50, $.50). Estimation of the likelihood function described in Equation A1 in the Appendix yielded a WTP estimate of $.75 (SE = .034) with a 95% confidence interval of $.68–$.82. Note that 64% of the respondents who were faced with an initial bid of $1.00 responded affirmatively. However, 42% of these respondents were unwilling to pay $1.50 for these songs. These numbers are consistent with the pricing at Apple's iTunes, which was launched approximately three months after the conclusion of this study (March 2003). In Studies 1 and 2, we use this pretested model to understand the impact of factors such as negative incentives, individual differences, and positive incentives on WTP, as well as more direct piracy measures.

Method

We conducted Study 1 immediately after the RIAA lawsuit against university students in September 2004, in which RIAA sued file sharers at 21 universities across the country in federal courts. The purpose of this study was to examine the impact of three manipulated variables (risk, embarrassment, and public versus private consumption) and one measured variable (Machiavellianism) on a consumer's WTP to download a song. The consumption setting (public versus private) was a between-subjects factor, and risk and embarrassment were within-subjects factors. We also varied the initial bid between subjects at the same five levels that were used for the estimation of the DBDC model.

We expected that the level of risk involved in choosing to download illegally, along with the potential embarrassment of getting caught, would affect participants' WTP for a legal download, as H_1 and H_6 suggest. Therefore, we manipulated three levels of risk by telling participants the probability of getting caught: 5% (low), 45% (moderate), or 100% (high). We also manipulated two levels of embarrassment by giving participants an estimated level of embarrassment that they would feel if they were caught on a scale from 1 to 10: 2 out of 10 (low) or 9 out of 10 (high). Because these factors were within subjects, each participant responded to one WTP bid set for each of the six scenarios. For example, in the low-risk, low-embarrassment scenario, they were told, "If you download the song for free, you have a very small probability (5%) of getting caught. If you get caught, it will cause you a level of embarrassment that you would rate a 2 (low) on a scale from 1 to 10." After providing their responses to the bids based on this scenario, participants repeated the task for the remaining five scenarios, in which we varied risk and embarrassment levels.

We also expected that participants would be willing to pay more for a song played in a public situation, such as a party, than for a song played privately in his or her own home or car, as proposed in H_7. Therefore, we manipulated the consumption setting by telling participants to "imagine you are creating a compilation CD of your favorite songs from your favorite musical artist (or band)" either to "listen to this CD in your car while commuting to school or work" (private condition) or to "play this CD at a party for a group of friends who share your musical interests" (public condition). This resulted in a 3 (risk of being caught: 5% versus 45% versus 100%) × 2 (level of embarrassment: 2 versus 9) × 2 (consumption setting: private versus public) × 5 (initial bid: .10 versus .25 versus .50 versus .75 versus 1.00) mixed experimental design.

Participants were 386 undergraduate students who completed the questionnaire in exchange for partial credit in a marketing course. Given the within-subjects design, this led to 2316 possible responses, 1925 (83%) of which were complete and consistent across the bid sets (i.e., the remaining 291 responses were either not completed or completed incorrectly). On the first page of the questionnaire, participants read a short news clip about the recent lawsuit in which 532 university students were sued by the RIAA for illegal downloading. On the next page, they were told to imagine creating a compilation CD for either private or public consumption. Then, they gave WTP estimates for each for the six scenarios. Next, they completed the 20-item Machiavellian scale (Christie and Geis 1970), which contained items such as "The biggest difference between most criminals and other people is that criminals are stupid enough to get caught" and "These days, almost everybody cheats (or lies) to some extent," measured on a scale from 1 ("strongly disagree") to 7 ("strongly agree"). Finally, they answered questions about their musical tastes, Internet habits, and demographics.

Results

To understand the impact of perceived risk, embarrassment, public/private consumption, and Machiavellianism on WTP, we estimated various models on the basis of the likelihood function in Equation A1 of the Appendix. In addition to the variables of interest, we included three covariates pertaining to respondents' involvement in music: (1) whether they attended music concerts on a regular basis (attend), (2) whether they played any musical instruments (play), and (3) whether they owned an MP3 player (MP3). Finally, we also controlled for gender. Table 1 summarizes the results.

WTP. We estimated the mean WTP for this sample at $.471 (SE = .017). This was less than the corresponding estimate in the pretest, primarily because we conducted Study 1 a year after the launch of iTunes, thus setting price expectations at $.99 (or lower) for a song.

Risk, embarrassment, and consumption setting. As we hypothesized, risk had a significant, positive impact on WTP. As the risk coefficient in Table 1 suggests, on average, a 10% increase in the probability of being caught increased WTP by $.03. Thus, this result provides evidence in support of H_1. In estimating the WTP for the three separate risk conditions, we found that the mean WTP for low-, medium-, and high-risk conditions was $.34, $.48, and $.62, respectively. These results suggest that though an increased risk of being caught increases the WTP significantly, even the certainty of being caught does not increase the mean WTP levels to the $.99 price point set by legal download sites such as iTunes. The impact of embarrassment was not significant, and thus H_6 was not supported by the data. However, we found a significant, negative impact of a private consumption setting on WTP, suggesting that public consumption of music increases the WTP for downloads. The magnitude of the private coefficient suggests that music intended for public consumption commands an average price premium of $.12, which provides strong support for H_7.

Machiavellianism. Our analysis of the Machiavellian scale suggests a high level of interitem reliability among all 20 items of the Machiavellian scale ($\alpha = .75$). We also estimated the reliability of the three subscales of the Machiavellian scale: general negativism ($\alpha = .71$), duplicity ($\alpha = .70$), and distrust of others ($\alpha = .45$). Because the subscale reliabilities were lower than that of the overall scale, we combined all 20 items into an overall Machiavellian score for this analysis. As Model 1 (Table 1) shows, the duplicity and negativism subscales had significant, negative impacts on WTP, but the distrust subscale had no effect. Model 2 presents the results of combining the items to create one Machiavellian scale (referred to as Mach 1). The coefficient of Mach 1 is significant and negative, and these results provide strong evidence in support of H_4; that is, a consumer's Machiavellian score has a negative effect on his or her WTP for a download.

Demographics and psychographics. The median age of our sample was 21 years, with a range of 18–48 years. Of the respondents, 73% were employed on a full- or part-time basis, and 86% had high-speed Internet access at home. A significant majority (86%) admitted that they had previously downloaded songs for free from sites such as Kazaa, Morpheus, and Grokster. When asked about the primary reason for downloading without paying, 37% of the sample responded that "I hate the fact that buying CDs involved paying for songs that I do not want," 42% claimed that "[i]t is easier and/or more convenient than buying a CD," and 10% mentioned that they used it to "sample songs prior to

TABLE 1

Study 1 Results: Impact of Risk, Embarrassment, Consumption Setting, and Machiavellianism on WTP

	Model 1		**Model 2**	
Variables	**Coefficient**	**SE**	**Coefficient**	**SE**
Constant	.567**	.120	.57**	.118
Duplicity	–.004*	.002	—	—
Negativism	–.006*	.002	—	—
Distrust	–.005	.004	—	—
Mach 1	—	—	–.005**	.001
Risk	.003**	.0004	.003**	.0004
Embarrass	.005	.004	.005	.004
Private	–.112**	.033	–.120**	.033
MP3	–.049	.034	–.049	.033
Gender	.177**	.034	.177**	.034
Attend	–.103**	.037	–.103**	.036
Play	–.150**	.036	–.151**	.036

*$p < .01$.
**$p < .001$.

purchasing them in the form of CDs." Furthermore, in response to the individual bids, 36% of the responses were in the (no, no) category; that is, they would not pay the amount stated in both the initial and the second (lower) bid. Of these, 75% said that they "would not pay anything at all," and more than half (56%) stated that they would "only download songs if they were available for free." These responses are a clear indication of piracy in that they represent conditions under which people have no interest in paying anything for obtaining a song.

We found that women had a higher WTP (an incremental $.18) than men, which is consistent with prior findings (e.g., Bhattacharjee, Gopal, and Sanders 2003). Moreover, consumers who played musical instruments and attended concerts regularly (i.e., music aficionados) had a lower WTP than those who did not. People who played musical instruments had an average WTP of $.15 below those who did not, whereas regular concertgoers had a WTP that was $.10 below those who did not attend concerts regularly. Finally, the ownership of MP3 players had no impact on WTP.

Why did we find a significant difference in WTP as a result of consumption setting (H_7) but not embarrassment (H_6)? One possible explanation is that participants had difficulty imagining the embarrassment resulting from the statistical probabilities given. Alternatively, they may have expressed a higher WTP for public consumption, not because of potential embarrassment but rather because of a higher demand for digital quality for music consumed publicly than for music consumed privately. In any case, potential embarrassment does not appear to provide the best explanation for differences in piracy potential. Therefore, in Study 2, we turn to the effects of positive incentives and risk tolerance and attempt to illuminate further the effect of negative incentives on pirating tendencies.

Study 2: The Effects of Negative Incentives, Positive Incentives, and OSL on WTP and Piracy Preference

We conducted Study 2 in mid-2005, approximately a year after the RIAA lawsuit against university students mentioned in Study 1. The primary goal of this study was to investigate whether a consumer's OSL determines his or her overall tendency to pirate music and his or her reaction to a negative incentive (i.e., risk) when deciding whether to engage in piracy. Because high-OSL consumers desire novel, varied, and complex experiences and are willing to take the risks necessary to achieve such experiences, we expected them to be more prone to piracy and more immune to risk. In addition, we estimated the effects of adding positive incentives (e.g., user-friendly Web site design or exciting features) on all participants' WTP. Finally, we added a more direct measure of tendency to pirate at the end of the questionnaire.

Pretest

A potential limitation of Study 1 is that we manipulated risk by simply telling participants to imagine a probability level of getting caught. A more realistic manipulation would be to have participants read a news article about the risks of illegal downloading that alters their risk perceptions. Therefore, for Study 2, we pretested three fictional news articles about a new wave of RIAA lawsuits targeting university students. The low-risk article stated that the RIAA recently sued 120 computer users, including 11 students at 3 universities. The moderate-risk article stated that the RIAA sued 750 computer users, including 125 students at 13 universities. The high-risk article stated that the RIAA sued more than 2500 computer users, including 25 students at the participants' own university. Immediately after participants read one of the three news articles, we measured their levels of perceived risk and likelihood of getting caught for engaging in music piracy. Furthermore, to determine which positive incentives they would be most likely to pay for, we asked, "Would offering the following features make you more likely to use the legal, pay site?" (1 = "would not influence my decision to pay," and 7 = "would make me much more likely to pay"), for the following features of a legal music downloading Web site: extensive music catalog, free samples, easy navigation, a recommendation engine that remembers preferences, and ease of burning compilations. We also asked, in free-response format, for other features that might increase participants' WTP for legal downloads.

Participants were 273 business undergraduate students who completed the pretest in exchange for partial course credit. The results indicated that reading the moderate-risk article resulted in a higher level of perceived risk (M = 3.96 versus 3.46 on a seven-point scale; $t(179) = 2.13$, $p < .01$) and a higher perceived probability of getting caught (M = 3.04 versus 2.37; $t(179) = 3.35$, $p < .005$) than reading the low-risk article. The perceived risk and probability of getting caught actually decreased slightly between the moderate-risk and the high-risk scenarios, but the differences between these two scenarios were not significant (M = 3.96 versus 3.85 for risk, not significant [n.s.]; M = 3.04 versus 2.75 for probability of getting caught, n.s.). Therefore, we used only the low-risk and moderate-risk scenarios as manipulations in Study 2.

Of the positive incentives we offered, a Web site that offered easy-to-burn compilations was the most likely to make participants want to pay for the legal site (M = 4.98 out of 7), followed by free samples (M = 4.88), an extensive music catalog (M = 4.60), ease of navigation of the site (M = 4.48), and a recommendation engine that remembers preferences (M = 4.21). For the free responses, many participants indicated that they would never pay for music downloads. The most popular write-in response for a feature that would indicate use of a legal site was a low price or discount. Other frequent responses included the following: good organization or file storage; other available downloads, such as ringtones or video; fast download times; security/privacy measures; high digital quality; access to rare or exclusive songs; and a flat membership fee. We used many of these free responses to formulate our questions about positive incentives in Study 2.

Method

In Study 2, we manipulated two levels of risk, using the low- and moderate-risk (now called high-risk) news articles from the pretest. We also varied the initial bid for the WTP estimation at the same five levels as in Study 1, which resulted in a 2 (perceived risk: low versus high) × 5 (initial bid: .10 versus .25 versus .50 versus .75 versus 1.00) between-subjects design. Three hundred fifty-nine respondents completed the questionnaire for partial credit in a marketing course. They read the low- or high-risk news article, rated their perceived risk and the probability of getting caught, and then gave their WTP estimates for downloading a legal song, using the DBDC model. They were told that the song would be part of a compilation CD, and the consumption setting was unspecified. Next, they responded to the question, "Would offering the following features make it more likely for you to use the legal, pay Web site?" on a scale from 1 ("would not influence my decision to pay") to 7 ("would make me much more likely to pay"), as in the pretest. The positive incentives offered were (1) an extensive music catalog; (2) free samples of the music; (3) easy navigation of the site; (4) a recommendation engine that remembers preferences; (5) easy-to-burn compilation CDs; (6) extremely fast downloading time; (7) rare recordings, exclusive artists, and/or new releases; (8) extremely high digital quality; (9) live concerts on the Web site; and (10) availability of additional downloads, such as ringtones or video. They subsequently completed the sensation-seeking scale (Zuckerman 1979), which is an instrument used to measure OSL (Raju 1980), and several questions about their music preferences, downloading experience, and demographics. One of these questions was an overall measure of tendency to pirate: "Are you more likely to download from a free site or a pay site?" (1 = "much more likely to pay," and 7 = "much more likely to download for free").

Results

Manipulation checks. Participants who read the high-risk news article demonstrated a marginally higher level of perceived risk (M = 3.81 versus 3.52; F(1, 344) = 2.86, $p <$ .10) and a significantly higher level of perceived probability of getting caught (M = 2.80 versus 2.50; F(1, 344) = 4.40, $p < .05$) than participants who read the low-risk news article. Therefore, it appears that participants viewed illegal downloading as a riskier proposition after reading the high-risk article than after reading the low-risk article.

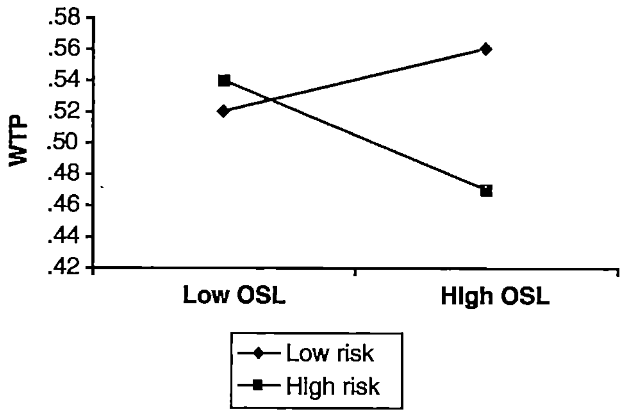
There was a high level of reliability among the 40 items on the sensation-seeking scale ($\alpha = .83$). Therefore, we combined the responses to all 40 items into an overall OSL score. When we split the participants into high- and low-OSL groups according to the median score, we found that high-OSL participants perceived illegal downloading as less risky than low-OSL participants (M = 3.37 versus 3.89; F(1, 269) = 18.01, $p < .0001$). High-OSL participants also believed that they would be less likely to get caught than low-OSL participants (M = 2.29 versus 2.92; F(1, 269) = 26.18, $p < .0001$).

Negative incentives. The estimated mean WTP in the high-risk condition was significantly higher (M = $.60) than that in the low-risk condition (M = $.43; t = 34.9, $p <$ .0001), providing strong support for H_1, which suggested that an increase in the perceived risk would lower the tendency to pirate (and thus increase WTP).

OSL. We found that the mean WTP for high-OSL participants was lower than the corresponding value for low-OSL participants (M = $.51 versus $.54; t = 3.92, $p <$ .0001). This supports H_2, which proposed an inverse relationship between OSL and the tendency to pirate. Furthermore, a correlation analysis of OSL with demographic and behavioral variables revealed that high-OSL consumers are more likely than low-OSL consumers to be men (r = .21, $p < .001$), to have a high-speed Internet connection (r = .24, $p < .0001$), and to attend concerts regularly (r = .20, $p <$ .001).

Negative incentives × OSL. H_3 suggested that OSL would moderate the relationship between perceived risk and WTP. To test this hypothesis, we estimated WTP for the four conditions of high risk/high OSL, high risk/low OSL, low risk/high OSL, and low risk/low OSL. The results (see Figure 1) provide support for the moderating effect of OSL proposed in H_3. When there was a low level of perceived risk, low-OSL participants had a lower WTP than high-OSL participants (M = $.52 versus $.56; t = 3.41, $p < .001$), but when there was a high level of perceived risk, low-OSL participants indicated a significantly higher WTP than high-OSL participants (M = $.54 versus $.47; t = 7.43, $p <$ .0001). Indeed, high-OSL participants significantly decreased their WTP when the perceived risk increased (from $.56 to $.47; t = 9.39, $p < .0001$). Therefore, although the RIAA's threats of potential arrest and/or lawsuits might halt piracy among low-OSL consumers, it is likely to backfire among high-OSL consumers and may even cause a higher rate of piracy.

FIGURE 1
Study 2 Results: The Effect of Risk and OSL on WTP (In $)

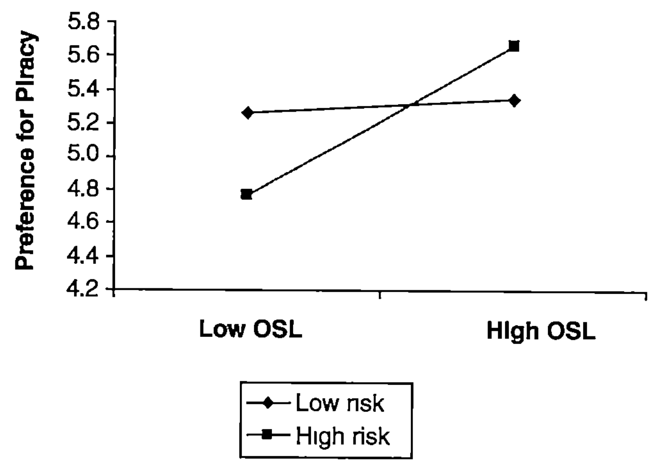


We also ran a 2 (risk: high versus low) × 2 (OSL: high versus low) analysis of variance (ANOVA) on participants' overall tendency to pirate (reported on a seven-point scale). In support of H_2, we found a significant main effect of OSL on tendency to pirate ($F(1, 340) = 4.92$, $p < .05$), such that high-OSL participants were more likely to engage in piracy than low-OSL participants ($M = 5.51$ versus 5.00). Furthermore, in support of H_3, we found a marginally significant risk × OSL interaction ($F(1, 340) = 3.21$, $p = .07$), which Figure 2 illustrates. When the risk level was low, high-OSL and low-OSL participants were equally likely to pirate ($M = 5.35$ versus 5.26; $F(1, 340) = .09$, n.s.), but when the risk level was high, high-OSL participants were significantly more likely to pirate than low-OSL participants ($M = 5.67$ versus 4.77; $F(1, 340) = 8.04$, $p < .005$). Indeed, high-OSL participants slightly (but not significantly) increased their likelihood to pirate when risk went up ($M = 5.35$ versus 5.67; $F(1, 340) = .82$, n.s.), whereas low-OSL participants marginally decreased their likelihood to pirate under the same circumstances ($M = 5.26$ versus 4.77; $F(1, 340) = 3.03$, $p = .08$).

Positive incentives. Respondents indicated a WTP of an additional amount for all the features that were proposed. The most popular features overall were fast downloading speed ($M = 5.66$ out of 7), availability of rare recordings ($M = 5.50$), high digital quality ($M = 5.47$), free samples ($M = 5.24$), and easy compilations (4.92).

Demographics and psychographics. Of the 344 participants, the mean age was 22.2 years ($SD = 3.68$), the minimum age was 19 years, and the maximum age was 51 years. The majority of participants worked part time (55.7%), and the remainder worked either full time (18.6%) or not at all (25.5%). Of the participants, 83% previously downloaded music for free from sites such as Kazaa, Morpheus, or Grokster. When asked the reasons for downloading free songs, the most popular response was, "I hate the fact that buying CDs involves paying for songs I do not want" (48.9%), followed by "It is easy to do and everyone does it anyway" (35.8%).

FIGURE 2
Study 2 Results: Preference for Piracy as a Function of Risk and OSL



Discussion

Although our results regarding positive incentives provide preliminary support for H_5, we cannot rule out the null hypothesis, because we did not measure WTP for each feature (or the tendency to pirate for each feature), and thus we cannot compare these measures with those for the unenhanced product. Therefore, in Study 3, we present hypothetical music retail Web sites with three levels of functionality (low, medium, and high) and compare the tendency to pirate among these three levels.

A possible alternative explanation raised in Study 2 is that our OSL findings are simply due to differences in perceived risk. The risk manipulation led to differences in the participants' perceived level of risk, which in turn led to differences in piracy and WTP measures. However, it is also possible that our moderating variable, OSL, simply measures differences in consumers' perceived risk rather than differences in willingness to take risks, as we intended. Indeed, high-OSL consumers reported lower levels of perceived risk than low-OSL consumers, as we reported previously. Furthermore, when we added perceived risk to the ANOVA model, it demonstrated a highly significant main effect on likelihood to pirate ($F(1, 336) = 20.04$, $p < .0001$), and the main effect of OSL and its interaction with manipulated risk were no longer significant. Participants who perceived a low level of risk were much more likely to pirate than those who perceived a high level of risk ($M = 5.78$ versus 4.66). In light of these findings, we cannot rule out the possible confound between OSL and perceived risk. Therefore, in Study 3, we use an alternative measure of consumer risk taking called the EXPL scale (Raju 1980). This scale measures a consumer's need for stimulation from the external environment and includes items such as "When I eat out, I like to try the most unusual items the restaurant serves, even if I am not sure I would like them." Although the EXPL scale is related to sensation seeking, it more specifically measures consumer risk taking, which is likely to take the form of variety seeking or early product adoption rather than general risk-seeking behavior. By using the EXPL scale to measure consumer risk taking, we hope to avoid confounding our manipulated variable (risk) and our measured variable (OSL).

Study 3: The Effects of Negative Incentives, Positive Incentives, and OSL on Preference for Piracy

We conducted Study 3 in fall 2006, approximately two years after the RIAA lawsuit against university students mentioned previously. Our main goal was to illuminate further the relationship among positive incentives (functionality), negative incentives (risk), and consumers' willingness to engage in music piracy. In addition, we further investigated the moderating role of OSL by using a more behavioral, consumer-oriented measure of individual differences in innovativeness, the EXPL scale (Raju 1980). We also measured piracy tendencies through both direct and indirect questioning (e.g., Fisher 1993).

Method

Study 3 used a 2 (risk: high versus low; between-subjects) × 3 (functionality: high versus medium versus low; within-subjects) × 2 (OSL: high versus low, between-subjects) mixed design. We manipulated risk by using the same news articles from the previous study. Furthermore, we described three different levels of functionality that the pay Web site offered and presented these to each participant as three separate scenarios. Scenario 1 described the low-functionality Web site, which was rated low in user friendliness by other users; offered songs that were low in digital quality and slow download speeds; and did not offer any samples, rare recordings, live concerts, or other downloads. Scenario 2 described the moderate-functionality Web site, which was rated high in user friendliness and offered free song samples, fast download speeds, and easy-to-burn compilations but did not offer any special items, such as rare recordings, live concerts, or other downloads. Scenario 3 described the high-functionality Web site, which offered all the features of the moderate-functionality Web site but also offered high-digital-quality downloads; an extensive catalog of rare recordings, exclusive artists, and new releases; viewing of live concerts; and additional files for download, such as ringtones and videos.

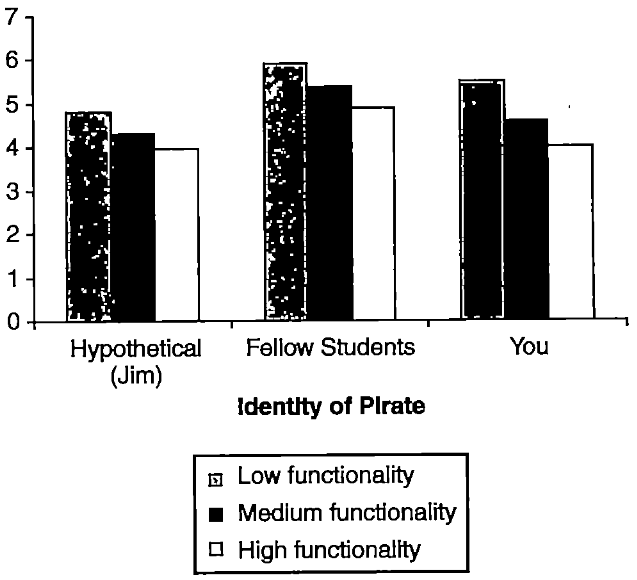
Participants were 165 undergraduate business students, who completed the questionnaire for partial credit in a marketing course. They first read the low- or high-risk news article and rated their perceived risk and the probability of getting caught. Next, for each of the three scenarios, they answered the following questions: (1) "Jim M., a student [at your university], regularly downloads music from the illegal Web site rather than paying for his music from the pay Web site described above. How acceptable do you find Jim's behavior?" (2) "How likely would other students [at your university] be to download from the pay Web site, versus the free, illegal Web site?" and (3) "How likely are you to download from the low-functionality pay Web site, versus the free, illegal Web site?" Because of the sensitive nature of these questions, we used a funneling technique; we first asked participants to judge the behavior of others and then to predict their own behavior. Each question appeared on a separate screen on the computer. We never used terms such as "piracy" or "theft." Instead, participants rated their own and others' likelihood of piracy on a seven-point scale, ranging from "much more likely to pay" (1) to "much more likely to download for free" (7). In the next section, they completed the 39-item EXPL scale (Raju 1980), an alternative measure of consumers' levels of OSL. Finally, they completed several demographic questions, such as age, gender, and income.

Results

Composite piracy measure. We analyzed the effects of risk and functionality on all three piracy measures (approval of Jim's piracy, likelihood of others' piracy, likelihood of own piracy) as a repeated measure, which we refer to as "identity." We found that risk (i.e., a negative incentive) had no significant effect on overall piracy (F(1, 161) = .59, n.s.), nor did it interact with exploratory behavior (F(1, 161) = .01, n.s.) or the identity of the pirate (F(2, 160) = .14, n.s.).

In support of H_5, the main effect of the pay Web site's functionality (i.e., a positive incentive) on overall piracy was highly significant (F(2, 162) = 56.82, $p < .0001$). Respondents were less likely to pirate music for themselves when functionality was moderate at the pay site than when functionality was low (M = 4.58 versus 5.49; F(1, 163) = 52.29, $p < .0001$), and they were even less likely to pirate when functionality was high than when it was moderate (M = 4.01 versus 4.58; F(1, 163) = 26.89, $p < .0001$). Moreover, they were less likely to expect others to pirate when functionality was high than when it was low (M = 4.87 versus 5.92; F(1, 163) = 34.81, $p < .0001$), and they were less approving of Jim's piracy when functionality was high than when it was low (M = 3.97 versus 4.81; F(1, 163) = 36.56, $p < .0001$). We also found a main effect of identity (F(2, 162) = 24.55, $p < .0001$), indicating that the participants believed it was more likely that other students would engage in piracy than that they would engage in piracy themselves. Finally, there was a significant functionality × identity interaction (F(4, 160) = 3.73, $p < .01$), suggesting that participants believed that differences in functionality would have stronger effects on their own piracy behavior than on the behavior of others. Figure 3 illustrates these findings.

FIGURE 3
Study 3 Results: Approval of Piracy/Likelihood to Pirate as a Result of Functionality of the Pay Site and Identity of the Pirate



OSL and a consumer's own likelihood to pirate. The 39 items on the EXPL scale (Raju 1980) were sufficiently intercorrelated (α = .83), but three of the scale items demonstrated negative correlations with the total. After we dropped these three items, Cronbach's alpha increased to .87. Therefore, we used the mean of the remaining 36 items as an overall measure of each participant's level of OSL.

Because participants predicted similar (and sometimes stronger) effects of the manipulated variables on their own behavior compared with others' behavior, we focused on this single dependent variable (own piracy behavior) for the remainder of our analysis. Therefore, we dropped identity and added OSL to the analysis, which resulted in a 2 (risk: high versus low) × 3 (functionality: high versus moderate versus low) × 2 (OSL: high versus low; median split) ANOVA. There was no main effect of risk on participants' predictions of their own piracy behavior. However, we found a three-way interaction among functionality, risk, and OSL ($F(2, 160) = 3.07, p < .05$). We show this interaction in Figure 4.

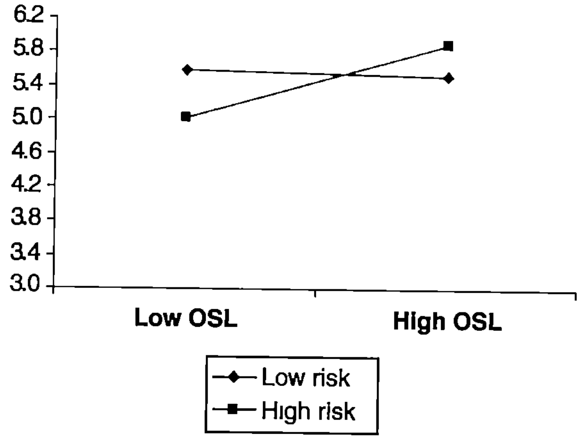
When the pay site was described as having low functionality, our results replicated those of Study 2 (in which participants likely assumed that the pay site did not offer any "extras," because none were indicated). In this condition, there was a marginally significant two-way interaction between risk and OSL ($F(1, 161) = 2.63, p = .10$). When the risk was low, participants were equally likely to pirate, regardless of whether they scored high or low on OSL (M = 5.54 versus 5.57; $F(1, 161) = .01$, n.s.). In contrast, when the risk was high, those who scored high on OSL were significantly more likely to pirate than those who scored low (M = 5.90 versus 5.02; $F(1, 161) = 4.91, p < .05$). This finding provides additional support for H_3, which predicted a risk × OSL interaction. However, when the pay site was described as having moderate or high functionality, there were no significant effects of risk, OSL, or their interaction (all comparisons were not significant). This three-way interaction represents an important boundary condition to our prior findings. When the functionality at the pay site was high (or moderate), the differences in piracy tendencies between high- and low-OSL consumers were no longer significant.

Furthermore, the EXPL scale appears to be a more valid measure of the factor that moderates the relationship among risk, functionality, and tendency to pirate, because unlike the sensation-seeking scale, it is not confounded with perceived risk. Indeed, participants who scored high on EXPL actually rated the perceived risk of piracy slightly higher than did those who scored low on the scale (M = 3.88 versus 3.36; $F(1, 161) = 4.63, p < .05$), but that did not stop them from pirating. If our findings could be explained as simply the result of differences in perceived risk, we might expect consumers who are low on EXPL to engage in more piracy than those who are high on EXPL, contrary to our actual results.

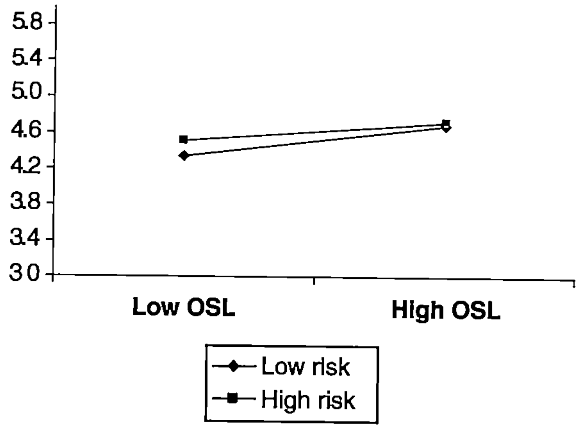
Demographics and psychographics. Of the 165 participants in this study, the mean age was 21.5 years (SD = 3.34), the minimum age was 19 years, and the maximum age was 45 years. Only 10% of participants worked full

FIGURE 4
Study 3 Results: Own Likelihood to Pirate as a Result of Functionality of the Pay Site, Risk, and OSL

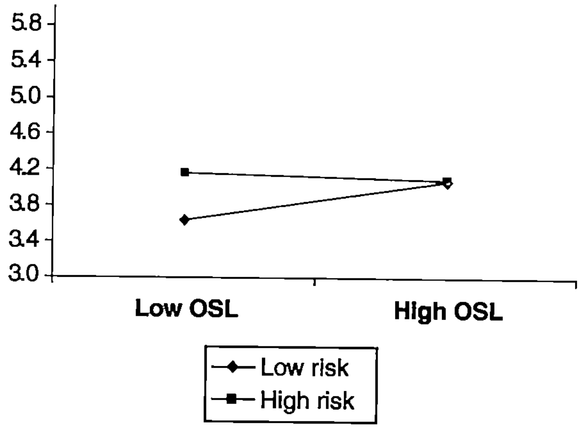
A: Legal Option Has Low Functionality



B: Legal Option Has Moderate Functionality



C: Legal Option Has High Functionality



time, 62% worked part time, and 28% did not work at all. When asked for their incomes, 59.4% of respondents indicated incomes of less than \$10,000 per year, 25.5% indicated \$10,000–\$24,999 per year, 9.7% indicated \$25,000–\$39,999 per year, and .5% indicated more than \$40,000 per year. When age and income were included in our model, they both demonstrated significant main effects on a consumer's own likelihood to pirate. Older participants were significantly less likely to pirate than younger participants ($F(1, 161) = 5.92$, $p < .05$), regardless of the level of risk or functionality. Moreover, higher-income participants were significantly less likely to pirate than lower-income participants ($F(1, 161) = 6.57$, $p = .01$), regardless of the level of risk or functionality. These findings suggest that a large proportion of piracy occurs within the segment of the population that is under the age of 21 and earns less than \$10,000 per year. A large percentage of participants had previously downloaded music for free from sites such as Kazaa, Morpheus, or Grokster (76.4%), and as in Study 2, the most popular reason for this piracy was "I hate the fact that buying CDs involves paying for songs I do not want" (46.0%). Therefore, we believe that our sample of undergraduate students adequately reflects the population that recording companies should target with their antipiracy strategies.

Discussion

The results of Study 3 offer further evidence that increasing the risk and/or consequences of illegal piracy is unlikely to discourage this undesirable behavior for certain consumers—namely, those with a high tolerance for risk. On the contrary, for these consumers, such a policy might actually backfire, causing them instead to increase their piracy levels. However, our results also provide some hope for recording companies; companies can decrease piracy behavior by offering their music on legal, pay Web sites that offer enhanced features, such as a user-friendly environment, live concerts, and downloadable ringtones. Indeed, when these upgraded features are present, potential pirates do not allow other variables, such as risk and exploratory behavior, to influence their piracy decision.

Although it may seem unsurprising that our effects were much stronger for the within-subjects factor (functionality) than for the between-subjects factor (perceived risk), we argue that our manipulations accurately reflect the real-world environment that potential music pirates encounter. Consumers are likely to estimate only one level of perceived risk before making the piracy decision, but they may allow the offerings of several retail music Web sites (e.g., iTunes, Real Rhapsody, Yahoo Music) to influence their decision.

General Discussion

The results of these studies indicate that the RIAA's efforts to use fear and shame to stop digital piracy may be effective only for certain segments of consumers. In Study 1, we found that a 10% increase in perceived probability of getting caught resulted in an increase of only \$.03 in consumers' WTP for a legal download. Furthermore, in Study 2, we demonstrated that for consumers with high levels of optimum stimulation (and, thus, higher tolerance for risk), increasing the perceived risk might actually backfire by slightly increasing their likelihood to pirate. Finally, the results of Study 3 indicate that whereas increasing the perceived risk may have a limited effect on decreasing piracy, offering enhanced Web site features can be highly effective. We found that all participants were less likely to pirate when the alternative, pay Web site offered features such as extensive music catalogs and the availability of extras, such as rare recordings, live concerts, and downloadable ringtones and videos.

We recommend that the RIAA make additional efforts to work with legal Web sites to make such features available. Some consumers have already demonstrated a WTP of up to \$2.49 for a partial song (e.g., a ringtone) downloaded directly to their cell phones, and they may be willing to pay as much or more to download a rare recording (Fry 2005). Furthermore, some consumers are paying anywhere from \$6.98 to \$13.98 to download live concerts (Smith 2005). Therefore, the music industry must continue to engage consumers with creative, engaging features to encourage legal downloading. We also agree with Bhatia, Gay, and Honey (2003), who suggest that when music distribution inevitably shifts to the Internet, music companies can cut costs by shifting much of the research and promotional costs associated with signing new artists to the Internet. For example, consumers willingly rate new songs when they are provided with a free download, and the subsequent online ratings charts may lead other consumers to pay to download music from these new artists. Furthermore, tiered pricing models might allow music companies to extract the maximum consumer surplus from each individual music track.

A limitation of the current research is that it included only university students (though the pretest sample included both graduate students and executives). On the one hand, this sample is highly relevant because university students represent a major target market for music companies. On the other hand, the WTP of university students might not accurately reflect the WTP of other populations of music customers. Optimal stimulation levels tend to decrease with age (Zuckerman 1979) and employment level (Raju 1980), and therefore older consumers might be less likely to pirate under all conditions and, in particular, when getting caught might result in professional embarrassment or even a job loss. Consistent with this notion, we found that even among the student sample, older participants and those with higher incomes were significantly less likely to pirate than other participants. Because college students are more likely than other demographic groups to pirate music, we believe that our sample is highly appropriate when studying the factors that may significantly decrease piracy levels. However, because of our use of a student sample, we are unable to generalize our findings to all populations. It is possible that a consumer's age interacts with our manipulated variables in unanticipated ways, and we recommend this avenue for further research.

A surprising finding was how little social norms against piracy seemed to influence participants' responses. The threat of embarrassment did nothing to increase their WTP for legal music (or, correspondingly, to decrease their intentions to pirate). Furthermore, we did not find significant differences between participants' own intentions to pirate and their projections of other college students' intentions to pirate (Study 3). These findings suggest that socially desirable responding was not a factor for the participants. Indeed, they did not appear to view music piracy as a transgression of social norms. If anything, digital piracy is the social norm among this segment of consumers. In support of this notion, only 5% of Generation Y consumers agreed that downloading music for free from the Internet was wrong or unethical (Freestone and Mitchell 2004). Therefore, attempts to curb piracy through appeals to consumers' sense of ethics are unlikely to be effective. Conversely, appeals to ethics and social norms might prove to be more effective among older, nonstudent populations.

Appendix
DBDC Model

To develop a framework for analyzing WTP, we define WTP* as a latent variable that represents the true WTP for downloading a song of choice. In other words, WTP* represents the amount of income forsaken by a consumer who is indifferent between a status quo state (e.g., maintaining an income of y by not purchasing a song or pirating it) and a new state (in which income is reduced to [y – WTP*] as a result of a legal purchase). To estimate the impact of covariates on WTP, we express WTP* as $WTP^* = x\beta + \varepsilon$, where x represents the vector of covariates and ε represents an i.i.d. error term with mean zero and standard deviation σ. Although WTP* cannot be observed, we can infer a range for the actual WTP by presenting respondents with a sequence of two bids and asking them whether their WTP equals or exceeds that bid (Haab and McConnell 2002; Mitchell and Carson 1989). The magnitude of the second bid depends on the answer (yes/no) to the first bid. With the initial bid denoted as B_I, respondents would be asked whether they would purchase the product if it were priced at B_I. If the answer is yes, the respondents are presented with a new bid B_H, where $B_H > B_I$. However, if the answer is no, they are presented with $B_L < B_I$. Different groups of respondents are presented with different bid sets based on a reasonable range established through prior research. Thus, the four outcomes may be represented as $D_{1i} = 1$ if $WTP_i < B_{Li}$ (no–no), $D_{2i} = 1$ if $B_{Li} \leq WTP_i < B_{Ii}$ (no–yes), $D_{3i} = 1$ if $B_{Ii} \leq WTP_i < B_{Hi}$ (yes–no), and $D_{4i} = 1$ if $B_{Hi} \leq WTP_i$ (yes–yes).

Respondents indexed by D_{1i} (those with a "no" response to both the initial bid B_{Ii} and the follow-up lower bid B_{Li}) represent consumers with a WTP between negative infinity and B_{Li}. Therefore, their contribution to the likelihood function can be written as

$$F\left(\frac{B_{Li} - x_i\beta}{\sigma}\right)^{D_{1i}},$$

where F(·) represents the standard normal cumulative distribution function. Similarly, respondents indexed by D_{2i} (a "no" response to the initial bid and a "yes" response to the follow-up lower bid B_{Li}) represent consumers with $B_{Li} \leq WTP_i < B_{Ii}$, whereas D_{3i} (a "yes" response to the initial bid and a "no" response to the follow-up higher bid B_{Hi}) represents $B_{Ii} \leq WTP_i < B_{Hi}$. Finally, respondents with a "yes" response to both bids represent consumers with $B_{Hi} \leq WTP_i$ (yes–yes). Consequently, the (log-) likelihood function for the data representing the contributions of respondents in each of these four groups can be written as

$$\text{(A1)}\quad LnL = \sum_{i=1}^{N}\left[D_{1i}\ln F\left(\frac{B_{Li} - x_i\beta}{\sigma}\right)\right.$$
$$+ D_{2i}\left\{\ln\left[F\left(\frac{B_{Ii} - x_i\beta}{\sigma}\right) - F\left(\frac{B_{Li} - x_i\beta}{\sigma}\right)\right]\right\}$$
$$+ D_{3i}\left\{\ln\left[F\left(\frac{B_{Hi} - x_i\beta}{\sigma}\right) - F\left(\frac{B_{Ii} - x_i\beta}{\sigma}\right)\right]\right\}$$
$$\left.+ D_{4i}\left\{\ln\left[1 - F\left(\frac{B_{Hi} - x_i\beta}{\sigma}\right)\right]\right\}\right].$$

Each of the four terms in Equation A1 represents the probability of a (no–no), (no–yes), (yes–no), and (yes–yes) response to each pair of bids, respectively. The estimated coefficients reflect the marginal impact of the respective covariates on the WTP for downloading a single song. The WTP for the song itself is estimated as $E(WTP) = (\bar{X}\hat{\beta})$, where $\bar{X}$ is a vector of means of the independent variables.

REFERENCES

Arrow, Kenneth, R. Solow, P.R. Portney, E.E. Leamer, R. Radner, and H. Schuman (1993), "Report of the NOAA Panel on Contingent Valuation," *Federal Register*, 58 (January), 4601–4614.

Bakos, Yannis and Erik Brynjolfsson (1999), "Bundling Information Goods: Pricing, Profits and Efficiency," *Management Science*, 45 (12), 1613–30.

——— and ——— (2000), "Bundling and Competition on the Internet," *Marketing Science*, 19 (1), 63–82.

Barker, J. Cam (2004), "Grossly Excessive Penalties in the Battle Against Illegal File-Sharing," *Texas Law Review*, 83 (2), 525–59.

Bauer, Raymond A. (1960), "Consumer Behavior as Risk Taking," in *Dynamic Marketing for a Changing World*, R.S. Hancock, ed. Chicago: American Marketing Association, 389–98.

BBC News (2007), "Apple Seeks Online Music Shake-Up," (February 7), (accessed August 20, 2007), [available at http://news.bbc.co.uk/2/hi/business/6337275.stm].

Bechtold, S.A. (2004), "Digital Rights Management in the United States and Europe," *American Journal of Comparative Law*, 52, 323–82.

Bhatia, G. Krishan, Richard C. Gay, and W. Ross Honey (2003), "Windows into the Future: How Lessons from Hollywood Will

Shape the Music Industry," *Journal of Interactive Marketing*, 17 (2), 70–80.

Bhattacharjee, Sudip, Ram D. Gopal, and G. Lawrence Sanders (2003), "Digital Music and Online Sharing: Software Piracy 2.0?" *Communications of the ACM*, 46 (7), 107–111.

Bishop, Richard C. and Thomas A. Heberlein (1979), "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics*, 61 (5), 926–30.

Christie, Richard and Florence L. Geis (1970), *Studies in Machiavellianism*. New York: Academic Press.

Connolly, Marie and Alan B. Krueger (2005), "Rockonomics: The Economics of Popular Music," in *Handbook of the Economics of Arts and Culture*, V.A. Ginsburgh and D. Throsby, eds. Amsterdam: North Holland, 667–720.

Dahl, Darren W., Rajesh V. Manchanda, and Jennifer J. Argo (2001), "Embarrassment in Consumer Purchase: The Roles of Social Presence and Purchase Familiarity," *Journal of Consumer Research*, 28 (December), 473–81.

Easley, Robert F., John G. Michel, and Sarv Devaraj (2003), "The MP3 Open Standard and the Music Industry's Response to Internet Piracy," *Communications of the ACM*, 46 (11), 91–96.

Fenigstein, Allan (1979), "Self-Consciousness, Self-Attention, and Social Interaction," *Journal of Personality and Social Psychology*, 37 (1), 75–86.

Fine, Michael (2000), "SoundScan Study on Napster Use and Loss of Sales," report by the chief executive officer of SoundScan, engaged by the plaintiffs in the action, *A&M Records Inc. et al. v. Napster, Inc.*, (accessed October 11, 2005), [available at http://news.corporate.findlaw.com/hdocs/docs/napster/riaa/fine.pdf].

Fisher, Robert J. (1993), "Social Desirability Bias and the Validity of Indirect Questioning," *Journal of Consumer Research*, 20 (September), 303–315.

Freestone, Oliver and Vincent-Wayne Mitchell (2004), "Generation Y Attitudes Towards E-Ethics and Internet-Related Misbehaviours," *Journal of Business Ethics*, 54, 121–28.

Fry, Jason (2005), "The Ringtone Riddle: How Ringtones' Success Has Led Labels to Verge of a Misstep on Music Pricing," *The Wall Street Journal*, (November 21), (accessed November 21, 2005), [available at http://online wsj.com/public/article/SB113199653761196701-dj6iAi9UQT4ti5gEDwwscCoaKX8_20061120.html?mod=blogs].

Haab, Timothy C. and Kenneth E. McConnell (2002), *Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation*. Northampton, MA: Edward Elgar Publishing.

Hansell, Saul (2005), "File-Sharing Services Seek Pact with Record Studios," *The New York Times*, (September 20), (accessed September 20, 2005), [available at http://www.nytimes.com/2005/09/20/technology/20grokster.html?ex=1284868800&en=1b7f64e6ecefa852&ei=5090&partner=rssuserland&emc=rss].

Jedidi, Kamal and Z. John Zhang (2002), "Augmenting Conjoint Analysis to Estimate Consumer Reservation Price," *Management Science*, 48 (10), 1360–68.

Johnson, Eric J., Steven Bellman, and Gerald L. Lohse (2003), "Cognitive Lock-In and the Power Law of Practice," *Journal of Marketing*, 67 (April), 62–75

Kahneman, Daniel and Amos Tversky (1979), "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263–91.

Keltner, Dacher and Brenda N. Buswell (1997), "Embarrassment: Its Distinct Form and Appeasement Functions," *Psychological Bulletin*, 122 (3), 250–70.

Leuba, Clarence (1955), "Toward Some Integration of Learning Theories: The Concept of Optimal Stimulation," *Psychological Reports*, 1, 27–33.

Levin, Aron M., Mary Conway Dato-on, and Kenneth Rhee (2004), "Money for Nothing and Hits for Free: The Ethics of Downloading Music from Peer to Peer Web Sites," *Journal of Marketing Theory and Practice*, 12 (1), 48–60.

Liebowitz, Stan J. (2004), "Will MP3 Downloads Annihilate the Record Industry? The Evidence so Far," *Advances in the Study of Entrepreneurship, Innovation, and Economic Growth*, 15, 229–60.

Lynch, John G., Jr., and Dan Ariely (2000), "Wine Online: Search Costs Affect Competition on Price, Quality, and Distribution," *Marketing Science*, 19 (1), 83–103.

Mandel, Naomi (2003), "Shifting Selves and Decision Making: The Effects of Self-Construal Priming on Risk Taking," *Journal of Consumer Research*, 30 (June), 30–40.

Mehrabian, Albert and James A. Russell (1974), *An Approach to Environmental Psychology*. Cambridge, MA: MIT Press.

Mick, David Glen (1996), "Are Studies of Dark Side Variables Confounded by Socially Desirable Responding? The Case of Materialism," *Journal of Consumer Research*, 23 (September), 106–119.

Miller, Rowland S. (1992), "The Nature and Severity of Self-Reported Embarrassing Circumstances," *Personality and Social Psychology Bulletin*, 18 (April), 190–98.

Mitchell Robert C. and Richard T. Carson (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.

Moe, Wendy W. and Peter S. Fader (2001), "Modeling Hedonic Portfolio Products: A Joint Segmentation Analysis of Music Compact Disc Sales," *Journal of Marketing Research*, 38 (August), 376–85.

Peitz, Martin and Patrick Waelbroeck (2004), "An Economist's Guide to Digital Music," CESifo Working Paper Series No. 1333, (November), (accessed August 20, 2007), [available at http://ssrn.com/abstract=628961].

Raghu T.S., Rajiv K. Sinha, Orneita Burton, and Ajay Vinze (2005), "Open Source Software and Piracy Controls: The Market Impact of Protection Strategies," Center for Advancement of Business with Information Technology working paper, W.P. Carey School of Business, Arizona State University.

Raju, P.S. (1980), "Optimum Stimulation Level: Its Relationship to Personality, Demographics, and Exploratory Behavior," *Journal of Consumer Research*, 7 (December), 272–82.

Reuters (2005), "Recording Industry Sues More U.S. File-Swappers," (accessed August 31, 2005), [available at http://news.com.com/RIAA+sues+754+more+file+swappers/2110-1027_3-5845136.html].

Schierman, Michael J. and Guy L. Rowland (1985), "Sensation Seeking and Selection of Entertainment," *Personality and Individual Differences*, 6 (5), 599–603.

Shapiro, Carl and Hal R. Varian (1999), *Information Rules*. Boston: Harvard Business School Press.

Smith, Ethan (2005), "Bootlegs Go Corporate: Record Companies Sell Instant Recordings of Concerts in Effort to Combat Piracy," *The Wall Street Journal*, (September 27), (accessed September 27, 2005), [available at http://online.wsj.com/article/SB112777091132852489.html].

Smith, Richard D., Jan A. Olsen, and Harris Anthony (1999), "A Review of Methodological Issues in the Conduct of WTP Studies in Health Care I: Construction and Specification of the Contingent Market CHPE," Working Paper No. 84, Center for Health Program Evaluation, Monash University.

Sullivan, Andy (2005), "Plugged In: 'Recommendation Engines' Lead Fans to Music," *Reuters*, (September 23), (accessed November 12, 2005), [available at http://findarticles.com/p/articles/mi_zdpcm/is_200509/ai_ni5615005/].

Wade, Jared (2004), "The Music Industry's War on Piracy," *Risk Management*, 51 (2), 10–15.

Weinstein, Eugene and Jerald Martin (1969), "Generality of Willingness to Take Risks," *Psychological Reports*, 24 (2), 499–501.

Wertenbroch, Klaus and Bernd Skiera (2002), "Measuring Consumers' Willingness to Pay at the Point of Purchase," *Journal of Marketing Research*, 39 (May), 228–41.

Zuckerman, Marvin (1979), *Sensation Seeking: Beyond the Optimal Level of Arousal*. Hillsdale, NJ: Lawrence Erlbaum Associates.

——— (1994), *Behavioral Expressions and Biosocial Bases of Sensation Seeking*. Cambridge, UK: Cambridge University Press.

Ramesh K.S. Rao & Neeraj Bharadwaj

Marketing Initiatives, Expected Cash Flows, and Shareholders' Wealth

It is widely accepted that marketing initiatives help firms acquire and retain customers. However, the link between the cash flows generated by customers' purchases and shareholders' wealth is not fully understood, and the literature is increasingly advocating the need for greater "marketing accountability." In this interdisciplinary research, the authors adapt the theory of firm valuation from finance and show that a marketing action can affect the shareholders' wealth by (1) determining the firm's net present value (the "stock price effect") and (2) potentially reducing the firm's cash needs (the "released working capital effect"). In demonstrating these results, the authors advance marketing theory by answering calls made in the literature to link marketing actions to the investors' cash flows, the firm's working capital needs, and the owners' wealth. The approach also shows that by reducing the firm's cash needs, marketing can increase the firm's productivity (operating efficiency). This imputes a critically important role for marketing. By increasing productivity, marketing can increase the firm's competitive posture and, thus, its long-term viability.

Keywords: marketing strategy, marketing accountability, firm value, shareholder wealth, productivity

There is increasing recognition in marketing that the field must become more financially accountable (e.g., Lehmann 2004; Rust, Zahorik, and Keiningham 1995; Srivastava, Shervani, and Fahey 1997). Because managers of a publicly traded firm have a fiduciary duty to maximize the shareholders' wealth, there is more discussion of how marketing actions must be aimed toward this goal. This view is implicit in the American Marketing Association's definition of marketing accountability: "The responsibility for the systematic management of marketing resources and processes to achieve measurable gains ... *while ... increasing the value of the corporation*" (emphasis added).[1] Thus, although a marketing initiative can lead to multiple meaningful outcomes (e.g., customer satisfaction, loyalty, market share), marketing accountability occurs only if these outcomes also ultimately serve to increase the stockholders' wealth.[2]

Firm value is the discounted value of the cash flows distributed by the firm to its securityholders. Therefore, it has been suggested that marketing activities must be linked to cash flows (Anderson 1982; Day and Fahey 1988; Doyle 2000; Rust et al. 2004). If the marketing activity requires an investment today, the key to justifying this investment is to articulate how it affects the investors' cash flows and the shareholders' wealth. To date, the literature has not adequately addressed this two-part question, and the answer is not obvious.

Scholars note the pressing need to formally trace how marketing can affect the investors' cash flows. For example, Srivastava, Shervani, and Fahey (1998, p. 15) conclude that the inclusion of cash flow as a focal variable in marketing research will "help marketers better understand the influence of marketing activities on shareholder value." Moreover, they suggest that marketing actions can lower the firm's working capital (cash) needs, but they acknowledge the need to investigate this intuition further. Stewart (2006) observes that it has become imperative for marketers to justify their actions in a manner similar to how managers elsewhere in the organization (e.g., finance) justify resource allocations (i.e., estimating cash flows under various economic scenarios). However, most marketers are not perceived as being able to deliver on marketing accountability.

In this regard, we embark on an interdisciplinary collaboration to link marketing actions to the investors' expected cash flows and the owners' wealth. The marketing literature has already demonstrated that a marketing action yields cognitive and affective outcomes, which in turn lead to purchase. Therefore, we do not explicitly discuss the stages in the customer's decision-making process from the time the prospect becomes aware of the product (as a result of the marketing action) until actual purchase. We take for granted that a purchase results from the marketing action and that when the firm makes a sale, the customer derives greater utility from this purchase than from other offerings that may be available in the marketplace.

---

[1] We direct the reader interested in further discussion to Ambler and Roberts (2006).

[2] Because of its potential to create wealth for the investors, we treat a marketing initiative (or a marketing action) as an investment. These tactical actions can include a banner advertisement, a loyalty-building program, a packaging change, a price change, a service improvement initiative, and so on.

Ramesh K.S. Rao is McDermott Professor of Banking and Finance, Department of Finance, McCombs School of Business, University of Texas at Austin (e-mail: ramesh.rao@mccombs.utexas.edu). Neeraj Bharadwaj is Visiting Assistant Professor of Marketing, Graduate School of Business, University of Chicago (e-mail: Neeraj.Bharadwaj@ChicagoGSB.edu). This version of the article has benefited from the feedback provided by the anonymous *JM* reviewers.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

We show that to evaluate the effect of a marketing action on the investors' expected cash flows, it is first necessary to understand how it affects the entire probability distribution of sales (not just the mean or mathematical expectation of future sales). This is important for at least two reasons. First, through its effect on the shape of the probability distribution of sales revenues, a marketing investment determines the firm's potential for operating losses and, thus, its working capital (cash) needs. We show that in this way, a marketing action can influence the firm's cash needs. Second, the cash requirement, in conjunction with the expectation of sales revenues, determines the investors' expected cash flows and the stock price. Formalizing this causal linkage between a marketing action and the investors' cash flows yields new and counterintuitive insights. For example, it reveals that a marketing action may have no effect on expected sales and yet affect the owners' expected cash flows. It also shows that an action can actually increase the firm's cash requirement. As another example, two marketing actions with an identical impact on expected sales revenues can create different amounts of shareholders' wealth.

Using a comprehensive numerical example, we illustrate the process by which a firm can identify the marketing action that is optimal for the stockholders. We show that the impact of a marketing activity on shareholders' wealth is the sum of two wealth effects: a "stock price effect" and a "released working capital effect." The former is the shareholders' wealth increase (net present value [NPV]) that is implicit in the stock price. In addition, to the extent that a marketing activity can lower the firm's cash needs, the firm can return any freed-up working capital to the investors. Investors can then reinvest the "released" working capital in investments outside the stock markets (e.g., private equity, real estate, art, proprietorship businesses, foreign exchange) and create additional wealth. This is the released working capital effect.

We organize the rest of the article as follows: We begin by tracing the progression of marketing thought on the link between marketing actions and firm performance. Next, we establish the relationship between a marketing action and the firm's expected cash flows and then formalize how the initiative affects the shareholders' wealth. Subsequently, we illustrate our approach by comparing and contrasting several competing marketing initiatives in terms of their ability to create shareholder wealth. We conclude with a discussion of the implications and limitations of this research, and we explain how this article provides the necessary theoretical scaffolding for a promising research agenda.

Progression in Marketing Thought Regarding the Link Between Marketing Actions and Firm Performance

A central tenet in marketing is that the firm's marketing actions must satisfy customer requirements, with adequate consideration devoted to the costs incurred in achieving the firm's profit objective (e.g., Kotler and Levy 1969). Although it has been advocated that researchers move beyond profits and link marketing actions to firm value (Anderson 1982; Day and Fahey 1988; Srivastava, Shervani, and Fahey 1997), what is missing is a means by which to formally trace the effects of a marketing action to firm value (stock price).

A reason for this knowledge gap is that marketers have historically investigated how their activities affect a consumer's cognition, attitude, and/or behavior (for a review, see Gupta and Zeithaml 2006). However, the focus on metrics related to customer cognition, attitudes, and behavior (e.g., awareness of product, satisfaction, purchases) has prevented a full appreciation of the benefits accruing to the selling firm from customers' purchases (Hunt 2002). Because the firm's fiduciary duty is to its shareholders, linking marketing actions to customer-perceived metrics stops short of establishing accountability (Lehmann 2004).[3]

There has recently been an increasing discussion in the literature that marketing actions must be linked to cash flows because the cash flows distributed to investors ultimately determine firm value (stock price) and the wealth created for shareholders (Rappaport 1986). Some have gone as far as to assert that "at the end of the day, marketing is the creation of cash flows" (Ambler and Roberts 2006, p. 4). Authors have posited various links among marketing actions, cash flows, and firm performance. For example, Srivastava, Shervani, and Fahey (1998) suggest that marketing activities create brand equity, customer equity, channel equity, and/or customer satisfaction that enhance and accelerate cash flows and perhaps lower the firm's working capital (cash) needs. Rust and colleagues (2004) note that reducing or maintaining marketing actions to achieve a given or higher level of sales can augment the firm's cash flows. They also offer a comprehensive chain-of-effects model that traces how a firm's marketing initiatives affect the firm's overall condition. Others have documented a correlation between key metrics that result from marketing actions (i.e., customer satisfaction) and stock price (Anderson, Fornell, and Mazvancheryl 2004; Fornell et al. 2006; Gruca and Rego 2005).

The literature has begun to discuss how cash flows generated by a firm's customer portfolio can augment shareholder wealth (e.g., Gupta and Lehmann 2005; Rust, Lemon, and Zeithaml 2004). However, the implicit assumption in the field appears to be that a marketing action will increase sales, which in turn will increase expected cash flow, which in turn will increase stock price and, thus, shareholder wealth. This intuition is not necessarily valid.

---

[3] This is mandated by legislation at the federal and state levels. For example, the Minnesota statute passed in June 1987 (Minn. Stat. 302A.251 [5] [Supp. 1988]) allows a firm, "in considering the best interests of the corporation, to consider the interests of the corporation's employees, customers, suppliers, and creditors, the economy of the state and the nation, community, and societal considerations" Although this statute explicitly acknowledges that various stakeholders can influence firm activities, it holds that the firm's directors may consider the interests of the other parties only as long as they are in the best interests of the corporation (i.e., stockholders).

To establish why, we need to analytically formalize the link between the probability distribution of sales associated with a marketing action and shareholder wealth.

The Effect of Marketing Initiatives on the Shareholders' Expected Cash Flows and Wealth

Ambler and Roberts (2006, p. 4) contend that "alternative scenarios and plans can be compared ... using future cash flows," and they imply that unless marketers have a way to address the uncertainty associated with future outcomes, their use of discounted cash flow methods is flawed. This means that to link marketing to firm value, marketers must be able to address how their actions affect expected cash flows (the mathematical expectation of the probability distribution of future cash flows). This requires a modeling of the nature of the uncertainty facing the firm.

By definition, the expected cash flows depend on alternative cash flow realizations. In this context, there is evidence that buyers respond differently in different "states of the economy." For example, Lamey and colleagues (2007) find that private-label product purchases are asymmetric. They find that buying behavior depends on the state of the economy (i.e., their response to a marketing action is markedly different when the economy is flourishing versus when it is experiencing a downturn). They also suggest that this evidence of "cyclicality" in buying behavior has a direct bearing on the success of a marketing action. To recognize how this affects expected cash flows, it is necessary to model the nature of the economic uncertainty facing the marketer. Accordingly, we formalize the linkage among a marketing action, its impact on investors' cash flows in various economic states, and the firm's stock price.

Marketing Initiatives and Shareholders' Wealth

For simplicity, consider a publicly traded firm that produces just one offering, which may be a tangible good and/or a service. To abstract from issues of financial policy and complexities arising from intertemporal decision making, we assume that the firm is all equity (no debt) in a single-period world; we denote time as t_0 (today) and t_1 (next period). To focus on the key economic issues, we ignore risk aversion because it does not change our arguments; it changes only magnitudes in our numerical examples.

At t_0, the firm must make an investment of I in physical assets (e.g., plant and equipment) to produce the offering that is to be sold at t_1. In addition, the firm makes an investment, M, today in a marketing initiative to generate initial awareness and subsequent demand for the offering next period. The firm relies on a direct go-to-market (distribution) strategy. The firm also incurs operating expenses (e.g., rent, raw materials, labor) in an amount f (exogenously set in the markets at t_0) that are payable at t_1. For expositional ease, we use terminology from economics and refer to these providers of nonfinancial inputs (e.g., suppliers, employees, utility companies) as "factors." Unlike the stockholders, who hold market-traded securities, a preponderance of factor contracts are not traded (Grossman 1995). Therefore, we assume that the factors do not hold securities. We also assume that the firm operates in a competitive economy in which it is a price taker for its necessary operating inputs.

The firm's sales next period are determined jointly by exogenous realities (state of the economy) and by the firm's choice of marketing initiatives. The states-of-the-economy concept formalizes the notion that the marketing manager cannot be sure about the sales for his or her offering next period and that the sales are a random variable.

We assume that there are s states of the economy at t_1 and that the probability that each state will occur next period is π_s. Depending on the state of the economy at t_1, the firm can sell its offering at prices p_s. The quantity sales of its offering in state s is q_s.

The firm's marketing action determines how much it sells in each state. Its sales revenues in state s are $\tilde{R}_s = p_s q_s$, and the vector of revenues across all s states and the vector of associated state probabilities, π_s, together determine the probability distribution of sales revenues.[4]

The marketing initiative also determines the firm's operating income (revenue less operating expenses), $\tilde{X}_s$, in that state:

$$\tilde{X}_s = \tilde{R}_s - f = p_s q_s - f. \quad (1)$$

We assume that the firm has operating risk (it can have operating losses in one or more states). Thus, $p_s q_s < f$ for at least one state. This means that the firm's operating income in some states can be negative, and the factors cannot be paid in those states. The firm's maximum earnings shortfall is $\text{Max}_s(0, f - p_s q_s)$. This notation simply states that the maximum shortfall is the maximum across all s of the difference between operating expenses and revenues when the difference is positive (it is zero otherwise). Because the firm's marketing initiative determines q_s, it also determines the firm's maximum operating shortfall.

Although stockholders have limited liability and can walk away without covering the firm's operating expenses, it would be meaningless to illustrate the paths to shareholder wealth wherein it is implicitly assumed, *ex ante*, that the firm can avoid paying the factors in some states of the economy. It is important to preclude this possibility in the shareholder wealth theory. The difficulty is that unlike the securityholders, the factors do not hold securities. However, as we discuss in the Appendix, it is optimal for the firm to protect the factors by maintaining adequate working capital (cash) on its balance sheet at t_0. The cash serves to optimally "internalize" the firm's operating risk and, thus, to protect the factors.

How much working capital should the firm hold? At t_0, it is sufficient for the firm to retain cash in an amount L that is equal to the present value of the potential shortfall in the event of an economic downstate at t_1, where the present value is calculated using the riskless interest rate (the interest rate on default-free U.S. treasury obligations), r_f. Thus, the firm's working capital requirement at t_0 is as follows:

---

[4] Note that by determining q_s, the marketing initiative influences the probability distribution of sales revenues. The marketing initiative does not affect the probability of the state of the economy next period, π_s. However, because a probability distribution is, by definition, a listing of the various potential outcomes and their associated probabilities, it is correct to say that the choice of marketing initiative determines the probability distribution of sales revenues.

$$(2)\qquad L = \frac{\underset{s}{\text{Max}}(0, f - p_s q_s)}{1 + r_f}.$$

First, note that L is the (discounted) present value of the cash shortfall next period (not the present value of the next period's expected shortfall, as it would be if the factors held securities). Second, to the extent that the marketing action M determines prices and quantities in state s, it also determines the firm's current working capital needs. Third, note that because $(f - p_s q_s) > 0$ for some states, the cash requirement in Equation 2 is positive. Because L earns an interest rate r_f, its value next period is $L(1 + r_f)$. Because the cash holdings are set to be the present value of the maximum potential shortfall next period, the factors can now be paid even in states in which the firm has an operating shortfall. Thus, our approach ensures that the welfare not just of the stockholders but also of the factors (e.g., employees) is protected, *ex ante*.

The cash requirement implies that the firm's total t_0 investment, K, has three components: the investment in plant, equipment, and human capital (I); the investment in the marketing initiative (M); and the necessary cash (L):

$$(3)\qquad K = I + M + L.$$

Note from Equation 3 that the choice of marketing initiative determines the firm's capital requirements first through the cost of the initiative, M, and then through the firm's working capital needs, L, which, as Equation 2 shows, depends on how the marketing initiative affects the probability distribution of sales revenues.

A firm can allocate its given marketing budget M to alternative marketing actions. Thus, a decision that marketing managers often must make is to which initiative the firm should allocate its limited resources. To adopt the optimal initiative, the firm must understand how its choices affect the shareholders' wealth. A prerequisite to this understanding is knowledge of how the initiative affects the shareholders' expected cash flows.

Marketing Initiatives and Expected Cash Flows

Consider the t_1 expected cash flows to the shareholders, E(C). The firm has operating earnings, $\tilde{X}_s$, as in Equation 1. In addition, because its cash holdings, L, earn an interest rate r_f over this period, it is worth $L(1 + r_f)$ at t_1. Thus, in state s, the investors' cash flow is as follows:

$$(4)\qquad C_s = p_s q_s - f + L(1 + r_f).$$

Because the state s is unknown at t_0, the expected cash flow to shareholders (with expectations taken at t_0) is as follows:

$$(5)\qquad E(C) = \sum_s \pi_s[p_s q_s - f + L(1 + r_f)].$$

Equation 2 shows that by affecting the probability distribution of sales (q_s and/or p_s), the marketing investment can determine the firm's working capital needs, L. Equation 5 shows that in conjunction with the riskless interest rate, L also affects the investors' expected cash flows. Thus, Equation 5 implies the following:

> P_1: A marketing initiative affects investors' expected cash flows (a) directly, by determining the probability distribution of future sales revenues, and (b) indirectly, by determining the firm's current working capital needs.

Note that P_1 is a statement about cash flows. It does not indicate how the stockholders' wealth is affected by the marketing investment. We turn to this issue next.

Marketing Initiatives and Two Wealth Effects

In the subsequent discussion, we show that there are two marketing wealth effects to consider. For expositional convenience, we first discuss the stock price effect and illustrate it with an example. We then generalize the analysis to bring in the released working capital effect and to illustrate the combined effects numerically.

The stock price effect. A marketing action can have an immediate impact on the firm's stock price. The firm's stock price at t_0 is the expected cash flows to stockholders (Equation 5) discounted at the appropriate rate. This means that all the firm's future cash flows are discounted in the markets at the rate r_f. From Equation 5, the firm's stock price, S, at t_0 is as follows:

$$(6)\qquad S = \frac{E(C)}{1 + r_f} = \frac{\sum_s \pi_s[p_s q_s - f + L(1 + r_f)]}{1 + r_f}.$$

We now define NPV^I, the NPV of the investment in the firm as the present value of expected cash flows to investors (i.e., the stock price) less their initial investment, K:

$$(7)\qquad NPV^I = S - K = \frac{\sum_s \pi_s[p_s q_s - f + L(1 + r_f)]}{1 + r_f} - (I + M + L).$$

Because NPV^I is implicit in the stock price, we call this the stock price effect.

An important feature of firms operating under perfect competition is that it is insufficient simply to have the ability to create shareholder wealth. Firms must also be efficient in the wealth creation process. This efficiency is what we call productivity—the wealth created per dollar of capital invested. Firms that do not operate productively are destined to be displaced by the competition in the long run.

Productivity has several meanings in the economics literature (e.g., labor productivity, total factor productivity). In the marketing literature, Anderson, Fornell, and Rust (1997, p. 131) note that productivity refers to "the efficiency with which inputs are transformed into outputs." Here, we construe productivity in terms of the efficiency of the wealth creation process, so that the relevant output is the wealth the firm creates. The productivity of the wealth creation process in the firm, which we denote as η, measures how efficiently the firm uses its invested capital in creating NPV^I:

$$(8)\qquad \eta = \frac{NPV^I}{K}.$$

For simplicity, we assume that the firm has a fixed budget, M, and seeks to invest the entire budget in one of the three initiatives (A, B, or C). For all these initiatives, the firm's necessary investment to produce the offering, I, is \$3; the marketing budget available to stimulate demand, M, is \$2; the firm's operating expenses, f, are \$40; and the risk-free interest rate in the economy, r_f, is 10%. We also assume that there are two possible states of the economy next period, u (upstate) and d (downstate), with probabilities π_u and π_d, respectively. The numerical values of these probabilities, the prices at which the offering can be sold in each state, and the sales that each initiative will generate in each state appear in Table 1, Panel A.

Note that each initiative has asymmetric sales in the two states, meaning that the action has differential effects during a period of expansion (upstate) versus contraction (downstate) (see Lamey et al. 2007). For the immediate discussion, we focus on Initiative A, which we label the "base case." The information in Table 1, Panel A, yields the probability distribution of the firm's sales revenues (see also Figure 1, Panel A). The expected revenues are \$68. Given the probability distribution, the manager knows that the firm's maximum operating income shortfall is \$20 (\$20 revenue in state d less \$40 operating expense, f). From Equation 2, this implies that the firm must hold \$18.18 in cash today (the present value of the \$20 discounted at the 10% riskless rate) so that the factors can be paid in the event of an economic downstate. Thus, the firm's total investment from Equation 3, K, is \$23.18. With this knowledge, we can calculate the owner's expected cash flows next period as \$48 (see Figure 1, Panel B).

The firm's stock price, S, determined from Equation 6, is \$43.64. The total wealth created for the investors, NPV^I, is \$20.46. This is the stock price effect. From Equation 8, the firm's productivity, η, is .88. We summarize this base case, Initiative A, in Table 1, Panel B. Note that in this example, all the wealth increase to the shareholders is from NPV^I. That is, all the wealth increase accrues through the stock price effect.

The released working capital effect. To understand this effect, it is necessary to compare two marketing initiatives. Consider the case in which the firm's cash needs with a particular initiative are L (e.g., the base case, Initiative A). Now assume that the firm decides to allocate funds to the initiative that lowers L. This means that the second initiative "releases" working capital.

Let the cash released by this marketing activity be ΔL (the difference between the original and the new cash requirement). Because the firm does not need to retain this extra cash, it can return the released cash to the owners, who in turn can (i.e., at t_0) invest the released cash elsewhere in the economy in nonsecurity investments (e.g., real estate, gold, art).

We denote the wealth created by investment opportunities outside the securities markets as NPV^O. Let these investment opportunities have a productivity θ. This means that the outside wealth created, NPV^O, per dollar of investment in nonsecurity investments is θ. (The interpretation is the same as that for η, except that, now, the productivity is for cash invested outside the securities markets: $\theta = NPV^O/L$.)

The ability of an initiative to release cash that investors can reinvest outside the firm enables them to increase their wealth. We denote the wealth increase from this released cash as ΔW:

$$\Delta W = (\theta)\Delta L. \tag{9}$$

Because each dollar of investment outside the securities markets increases the shareholders' wealth by θ, the total wealth increase from the released cash, ΔL, is simply $(\theta)\Delta L$. This is the released working capital effect.[5] The working capital effect is positive as long as $\theta > 0$.

The total effect on the shareholders' wealth, denoted as Ω, is the sum of the two wealth effects of marketing: the NPV^I generated by the marketing initiative (Equation 7) and the wealth increase to shareholders from any released cash. That is,

$$\Omega = NPV^I + (\theta)\Delta L. \tag{10}$$

The next proposition summarizes the preceding discussion:

> P_2: A marketing investment can affect the shareholders' wealth in two ways: (a) by affecting their wealth as reflected in the stock price (the stock price effect) and (b) by affecting the wealth that investors can create outside the stock markets with any freed-up working capital (the released working capital effect).

If a marketing initiative is to benefit the shareholders, Ω must be positive. It is useful to note that simply increasing the investors' cash flows (in P_1) will not necessarily translate into increases in shareholders' wealth (in P_2). Indeed, under the Initiatives B and C, the shareholders' wealth increases even though the marketing initiatives lower the investors' expected cash flows.

We now examine Initiatives B and C (Table 1, Panel A) and compare them with the base case (Initiative A). The results of this analysis appear in Table 1, Panel B.

Table 1, Panel B, replicates the preceding analysis for Initiative B. The firm notices that the NPV^I under Initiative B (\$20.45) is slightly lower than the base case (\$20.46), but there is an important difference. Because the nature of the sales asymmetry in the upstate and downstate is different under Initiative B (see Table 1), this initiative requires the firm to hold less cash, L, (\$4.55) than Initiative A (\$18.18). Thus, Initiative B releases \$13.63 in cash (ΔL) compared with Initiative A. If this released cash can be invested by stockholders outside the firm at productivity $\theta = .9$, the released working capital effect under Initiative B, ΔW from Equation 9, is \$12.27. Together, the released working capital effect and the stock price effect yield a total wealth increase of \$32.72 to the firm's owners. After considering these joint effects of the initiative, the firm concludes that Initiative B is superior to Initiative A because it increases the stockholders' wealth by a larger amount (\$32.72 versus \$20.46).

---

[5] The released working capital effect disappears if the restrictive assumption is made that the released cash cannot be invested outside the securities markets. The Appendix elaborates on this.

TABLE 1
An Illustration of the Shareholder Wealth Implications of Marketing Initiatives

A. Details Regarding the Three Marketing Initiatives

Economic States	Probability of State	Prices	Quantity Generated by Marketing Initiative (in Units)		
			A	B	C
Upstate (u)	$\pi_u = .6$	$p_u = \$10$	$q_u = 10$	$q_u = 9$	$q_u = 8$
Downstate (d)	$\pi_d = .4$	$p_d = \$5$	$q_d = 4$	$q_d = 7$	$q_d = 6$

B. Stock Price and Released Working Capital Effects Under Alternative Marketing Initiatives

Economic Variable	Equation Number	Brief Description	Initiative A	Initiative B	Initiative C
		Sales revenue in upstate	\$100.00	\$90.00	\$80.00
		Sales revenue in downstate	20.00	35.00	30.00
$E(S) = \sum_s \pi_s \tilde{R}_s$		Expected sales revenues	68.00	68.00	60.00
$\tilde{X}_s$	1	Operating income (sales revenue less operating expenses)	28.00	28.00	20.00
L	2	Necessary cash (working capital) the firm must hold at t_0 so that it can cover payments to its factors (f) at t_1 in the event of an economic downstate.	18.18	4.55	9.09
K	3	The total capital required to generate cash flows.	23.18	9.55	14.09
E(C)	5	Expected cash flows to shareholders at t_1.	48.00	33.00	30.00
S	6	The firm's stock price at t_0 is the expected cash flows to stockholders discounted at the appropriate interest rate.	43.64	30.00	27.27
NPV^I	7	The wealth created for shareholders from selling the firm's market offering (the stock price effect).	20 46	20.45	13.18
η	8	The wealth created in the firm per dollar of invested capital.	.88	2.14	.94
ΔL		The cash released to investors at t_0 that they can invest elsewhere. It is incremental with respect to Marketing Initiative A.	Not applicable (this is base case)	\$18.18 − 4.55 \$13.63	\$18.18 − 9.09 \$ 9.09
ΔW	9	Extra wealth created for investors by cash released by marketing initiative. Assume returns that can be provided elsewhere, θ, are .9 (the released working capital effect).	No cash released	\$13.63 × .90 \$12.27	\$9.09 × .90 \$8.18
Ω	10	Total wealth increase is the sum of the stock price effect and the working capital effect.	\$20.46	\$20.45 +12.27 \$32.72	\$13.18 + 8.18 \$21.36

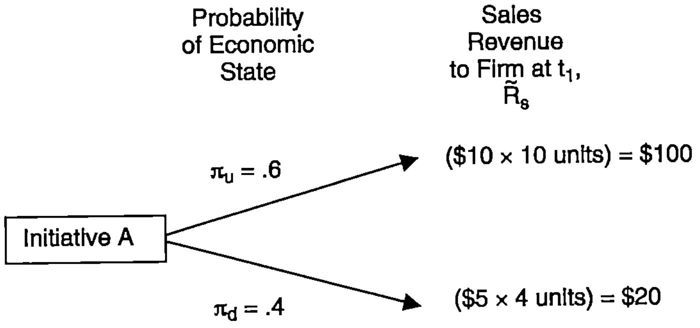
Notes: The stock price effect, NPV^I, is from Equation 7, and the released working capital effect, ΔW, is from Equation 9. The total wealth increase from shareholders, Ω, is the sum of the two effects, $NPV^I + \Delta W$

Initiative C is also better than Initiative A because it increases shareholder wealth by \$21.36 (versus the base case: \$20.46). In evaluating Initiative B in relation to Initiative C, the firm observes that the cash requirement, L from Equation 2, is higher for the latter case, thus making the capital requirement, K, higher and the productivity, η, lower than for Initiative B. On the basis of the overall effect on the shareholders' wealth, the firm concludes that Initiative B is

KP2722

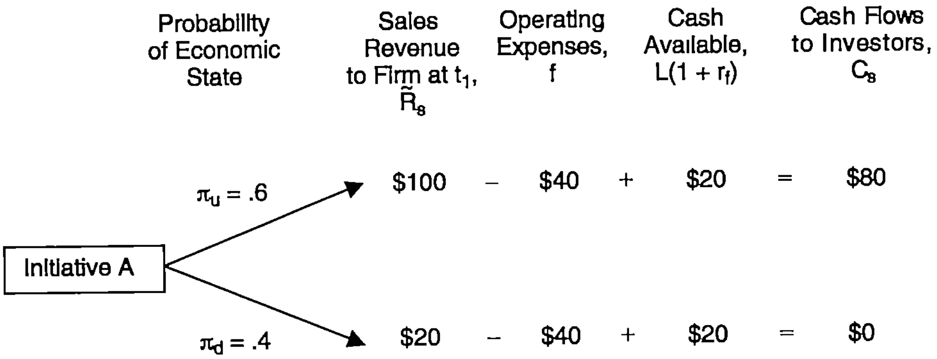
FIGURE 1
An Illustration of the Expected Cash Flow Implications of a Marketing Initiative

A. Expected Sales Revenues



Expected Sales $= \sum_s \pi_s \tilde{R}_s = (.6 \times \$100 + .4 \times \$20) = \$68.$

B. Expected Cash Flows to Investors


Probability of Economic State
Sales Revenue to Firm at t_1, $\tilde{R}_s$
Operating Expenses, f
Cash Available, $L(1 + r_f)$
Cash Flows to Investors, C_s
$\pi_u = .6$
$100 – $40 + $20 = $80
Initiative A
$\pi_d = .4$
$20 – $40 + $20 = $0

Expected Cash Flows $= \sum_s \pi_s C_s = (.6 \times \$80 + .4 \times \$0) = \$48.$

the optimal marketing action. That is, Initiative B has the most "accountability" to the firm's owners.

Some economic intuitions. The conventional thinking is that marketing actions benefit the owners by increasing the firm's expected sales, which in turn increase the investors' expected cash flows, which in turn increase the owner's current wealth. As we have demonstrated, however, the linkage between sales and the stockholders' wealth is more complex.

By comparing the three initiatives in Table 1, Panel B, we can draw several insights. Two initiatives with identical expected sales revenues (A, B) need not imply the same expected cash flows to investors (Initiative A offers higher cash flows). This is because the investors' expected cash flows depend on the firm's working capital needs, L, which depends on the "downside risk" (operating loss). For example, Initiative B has less downside risk, and it reduces the firm's capital investment, K ($9.55 versus $23.18), thus increasing the firm's internal productivity (2.14 versus .88). This amount is different across the three initiatives in Table 1, Panel B, because the shape of the probability distribution varies across the competing alternatives. It is also useful to note that Initiative C, with lower expected sales revenues than Initiative A, actually creates more wealth for shareholders than Initiative A ($21.36 versus $20.46). These effects and interpretations are sensitive to magnitudes of the various effects, and thus a case-by-case analysis is required to evaluate alternative actions. In addition, because the riskless interest rate determines the firm's working capital need, it follows that the wealth implications of a marketing action can vary with changes in economic conditions (interest rates).

Customer Equity and Intermediate Marketing Assets

There is a burgeoning literature that discusses how marketing creates intermediate assets (e.g., customer equity, brand equity) that result in sales over time, not necessarily immediately (see, e.g., Rust et al. 2004). Gupta, Lehman, and Stuart (2004) suggest that customer equity is "a useful proxy" for the firm's market value. In this regard, Rust, Lemon, and Zeithaml (2004) propose that firms can assess the effects of marketing expenditures on their customer equity, which they define as the total of the discounted lifetime value of cash flows summed over all the firm's current and potential customers.

Our approach to evaluating the impact of an action is for a single period in which a sale at t_1 results from an investment made at t_0. To integrate the customer equity notion into the single-period model of this article, it is necessary to redefine the revenues, $\tilde{R}_s$ in Equation 1, to include the value of customer assets. Thus, $\tilde{R}_s = p_s q_s + CE_s$. This means that the firm's working capital requirement in Equation 2 is also modified to include the value of the customer equity. This will lower the firm's working capital needs and boost the firm's productivity.

A more useful approach would extend our analysis to a multiperiod framework in which customer behavior is endogenous (i.e., buying behavior is explicitly modeled in the valuation analysis). For example, Rust, Lemon, and Zeithaml (2004) begin with a specification of individual-level customer utility, the behavior of individual customers, and customer-level brand switching to estimate the customer's lifetime value, which then determines customer equity (see their Equations 6 and 8). Such an approach to modeling customer buying behavior, when integrated with our shareholders' wealth approach, can also formalize how a marketing investment M affects the firm's revenues (i.e., the integration can potentially shed light on the functional relationship between M and $p_s q_s$). We believe that through a model of buying behavior, such an integration of customer equity and the firm value framework holds the promise of deeper insights into how marketing decisions affect the shareholders' wealth.

Discussion

The Boardroom Project estimates that marketing expenditures amount to 20%–25% of gross revenues at many firms (Stewart 2006).[6] However, marketers lack the means by which to link systematically the cash flows emanating from these actions to firm value. Therefore, researchers have called for the integration of the disciplines of marketing and finance (Day and Fahey 1988). Srivastava, Shervani, and Fahey (1998, p. 14) assert that "theories of marketing must be extended and broadened to include developments in finance, as indeed theories of finance must be extended and broadened to include recent developments in marketing."

This is precisely what we have tried to do in this article. We build on firm valuation theory in finance (Modigliani and Miller 1958), and by relaxing some traditional assumptions, we model the economic linkages among marketing actions, expected cash flows, and shareholders' wealth. This integration of marketing and finance yields important contributions.

The implicit assumption in the marketing literature is that a marketing initiative will increase sales, which in turn will increase expected cash flows, which in turn will increase stock price and shareholders' wealth. However, we demonstrate that the effect of marketing initiatives on expected cash flows and shareholders' wealth is not this straightforward.

First, we show that a marketing action affects the probability distribution of the firm's sales revenues (see our discussion preceding Equation 1). By recognizing that an action affects a probability distribution (rather than a point estimate), we answer several calls for research. Rust and colleagues (2004) note that further research is needed on the effects of the economic environment on firm value. Our states-of-the-world approach models alternative economic scenarios to be factored into an evaluation of the outcomes of marketing actions. It also accommodates exogenous economic factors (e.g., interest rates) into the marketer's decision-making process and thus contributes to the emerging literature on economic cyclicality (Lamey et al. 2007).

Second, we show that the probability distribution of sales revenues and the firm's operating details determine the firm's maximum operating shortfall and, thus, the firm's working capital (cash) needs (Equation 2). This reveals that two marketing initiatives with identical expected sales revenues can have differential working capital requirements (compare Initiatives A and B in Table 1, Panel B). Srivastava and Reibstein (2005) report that it is not well understood how marketing expenditures influence working capital and contend that it is rare to find articles that discuss the linkage. Our research is responsive to this recognized need.

Third, the cash requirement and the firm's probability distribution of operating income together determine the expected cash flows to the firm's owners (Equation 5). This reverses a standard intuition in marketing, as we noted previously. It implies that a marketing action that does not affect the firm's expected sales can still affect the investors' expected cash flows by altering the shape of the probability distribution of sales (compare Initiatives A and B in Table 1, Panel B).

Fourth, by affecting the firm's total capital requirements (Equation 3), a marketing initiative also affects the firm's productivity (Equation 8). We show that the firm's productivity can increase even if expected sales decrease (compare Initiatives A and C in Table1, Panel B). This extends the level of analysis in the prior studies of productivity (e.g., Anderson, Fornell, and Rust 1997; Rust et al. 2004) by examining the productivity of the investment.

Fifth, we establish that the implications of a marketing action on shareholder wealth are jointly determined by the stock price effect and the released working capital effect. By explicating the two paths by which marketing can augment shareholder wealth, our article contributes to making marketing more accountable (see Rust, Lemon, and Zeithaml 2004). A particularly nonobvious result is that even when an initiative does not increase expected sales, because of the released working capital effect, it can still be optimal for the firm (compare the total wealth effects under Initiatives A and C in Table 1, Panel B). By demonstrating that marketing initiatives can augment the firm's productivity (operating efficiency), we impute a new economic rationale for marketing. By increasing a firm's productivity, marketing can increase the firm's competitive posture and, thus, its long-term viability.

---

[6] The Boardroom Project is a group of leading marketing scientists and practitioners who, according to Stewart (2006), seek to "establish marketing measurement standards for continuous improvement of financial performance."

In regard to this article's implications for business practice, we note that the stature of marketing and its influence in the firm have been in reported decline (Webster, Malter, and Ganesan 2005). Traditional marketing activities, such as channel management and pricing, are being dispersed to other functional areas (Moorman and Rust 1999). Furthermore, few firms today have a chief marketing officer serving alongside a chief financial officer or a chief operating officer to influence the chief executive officer and the firm's corporate strategy (Kumar 2004). Even in firms that have a chief marketing officer, his or her presence is related neither positively nor negatively to firm performance (Nath and Mahajan 2008). In light of the heightened questioning regarding the importance of marketing, academics are increasingly asking how marketers can increase their accountability.

We reiterate that marketers must be able to demonstrate how their activities affect owners' wealth. This does not mean that customers are unimportant. It means that even as the firm addresses the needs of other constituencies, the manager's fiduciary duty is first and foremost to the firm's stockholders. Therefore, it is the responsibility of top management and the board of directors to reject initiatives that fail to increase the shareholders' wealth. In this context, the contribution of marketing actions must ultimately be assessed in light of their impact on the firm's primary goal (maximizing the shareholders' wealth).[7]

By framing their actions in accordance with the approach provided herein, marketers can justify their initiatives in the same way that financial managers make capital appropriation requests. As virtually every corporate finance textbook discusses (e.g., Brealey, Myers, and Allen 2006), financial managers present to the approving authority a capital budgeting analysis based on estimates of cash flows under various economic scenarios (e.g., optimistic, most likely, pessimistic). Then, they conduct a "sensitivity analysis"—an exercise that sheds light into which estimates are important or unimportant for ensuring a positive NPV outcome. Marketers can provide key inputs into this decision-making process by articulating how their actions can alter the probability distribution of sales (i.e., the sales revenues in each of the economic states of the world being examined) and by demonstrating that alternative marketing choices have different released working capital and stock price effects. In this sense, marketers are on the same playing field as their finance counterparts, and they too are subject to the same standards that are demanded of capital budgeting decisions. Indeed, the analytical approach we advocate herein makes marketing investment decisions de facto capital budgeting (investment) decisions.

We acknowledge that not all marketing phenomena can be adequately studied in a single-period model, and the need to extend our model to a multiperiod world is obvious. For example, the Marketing Science Institute and the Institute for Brand Science have recently suggested that an important question to explore is the extent to which changes in customer equity (as well as other marketing assets) affect firm value. This will require a longitudinal assessment of the relationship among marketing spending, customer equity, and stock price and would ideally be carried out for multiple firms in a given industry (as opposed to a single firm).

Further research can also usefully extend our analysis to a multiperiod world to capture persistence effects (e.g., Dekimpe and Hanssens 1995, 1999). This generalization will be beneficial for evaluating how heterogeneity in terms of payment timing (i.e., prepay versus postpay) affects NPV. Using a multiperiod model, researchers can also examine customer loyalty and churn rates as a decay function. However, it must be noted that discounting of multiperiod cash flows poses several technical challenges (Copeland and Weston 1979). It is our opinion that much is to be gained by further studying the single-period model before shifting attention to a multiperiod model.

In particular, it is necessary to study, in the single-period model, how marketing initiatives can affect the stockholders' risk. Srivastava and Reibstein (2005) suggest the possibility of developing a marketing-risk beta. Given that the notion of marketing risk is yet to be explicated, there exists a promising research opportunity to extend our analysis to address how marketing can affect the firm's financial-returns beta (systematic risk). In a shareholder wealth framework, the relevant risk is not variance but rather the covariance structure of asset returns (stock betas). Thus, the next step is to generalize our approach to model the implications of marketing on the stockholders' beta risk. Recent research has addressed this matter empirically (McAlister, Srinivasan, and Kim 2007), and our states-of-the-economy approach provides the foundation for studying how the firm, through its choice of marketing actions, can affect marketing assets (e.g., customer equity) and, in turn, the covariance of stock returns in light of exogenous economic variables. This would squarely link marketing to the risk-return framework of the firm valuation theory. The possibility that marketing can affect the firm's systematic risk and, thus, its cost of capital presents an exciting avenue for further research.

Appendix

This article shows that alternative marketing initiatives can have different implications for expected cash flows, for the firm's working capital (cash) needs, and for the firm's productivity. It also shows that marketing actions can affect the shareholders' wealth in two ways: the stock price effect and the released working capital effect. The purpose of this Appendix is to discuss (1) why working capital enters into our theory of shareholders' wealth and (2) why there are two (the stock price effect and the released working capital) wealth effects.

---

[7] For example, William Smithburg (1993), chief executive officer of Quaker Oats, offered the following: "The truth is, being successful at marketing or product development or manufacturing isn't enough. When your primary constituents, your shareholders, make an investment in your company, they want to build appreciation in the stock price and dividend. That's their criteria for buying it. And if chief executives don't understand what creates that movement, they won't get the job done."

The firm valuation theory in finance assumes that (1) the firm has no operating risk (zero probability of operating losses),[8] (2) all economic activity is in the securities markets (the "securities economy" assumption),[9] and (3) the firm can be valued as a perpetuity. With these assumptions, there is no role for cash in the theory, and the investors' expected cash flows are their expectations of the firm's net operating income (no cash implications).

However, numerous authors have criticized these assumptions as being too restrictive. In particular, there has been extensive discussion in finance and economics of the need to relax the securities economy assumption.[10]

Accordingly, we assume that the firm has operating risk and that there are two kinds of contracts—securities held by the providers of financial capital and nontraded factor contracts held by the owners of factors of production. To ensure that the welfare of the factors is protected in the theory, the firm must offer the factors, *ex ante*, their opportunity cost. An optimal way to ensure this requirement in the theory is for the firm to hold cash that is equal to the present value of the maximum shortfall (not the expected shortfall).[11] The cash requirement evaporates if the firm can be assumed to exist into perpetuity. The reason is that a perpetually viable firm can always overcome its cash shortfalls by issuing new securities.

Our more general assumptions rationalize working capital, and as Equation 5 shows, this redefines the investors' expected cash flows and, thus, the stock price. If a marketing decision can change the firm's working capital needs, the firm can return the cash to the owners, and the reduced capital requirement boosts the firm's productivity. In a securities economy, the only way that investors can reinvest this released cash is in the securities markets, but in well-functioning securities markets, all investments have zero NPV. Thus, the released working capital has no wealth effects for the shareholders. In contrast, in our analysis, we assume that the investors can reinvest the released cash in wealth, creating investments outside the stock markets. Thus, there is a second wealth effect to consider—the released working capital wealth effect—as Equation 9 shows.

---

[8] Modigliani and Miller's (1958) analysis assumes that the firm cannot have operating losses (i.e., they assume that operating income $\bar{X} > 0$) (see also Heins and Sprenkle 1969; Modigliani and Miller 1969, n. 2). Arrow and Hahn (1971, p. 21) also discuss this assumption.

[9] The standard theory assumes that the firm issues market-traded securities to providers of both nonfinancial and financial inputs. Thus, the theory assumes that *all* economic agents—labor, landlords, suppliers of processed goods, utility providers, and so on—are stockholders (for additional discussion of this assumption, see Cochrane 2001; Fama and Miller 1972; LeRoy and Werner 2001; Magill and Quinzii 1996)

[10] See Grossman's (1995) and Constantinides's (2002) presidential addresses. Similar ideas are expressed in Coase's (1992) Nobel address. See also Shiller's (1993, 2003) recent and highly influential works. LeRoy and Werner (2001) preface their book with a discussion of the restrictiveness of this assumption.

[11] There are several ways to rationalize the cash requirement. First, the factors cannot "bid up" the price of their inputs under the price-taker (perfect competition) assumption. Moreover, there are numerous exogenous restrictions (e.g , Robinson–Patman Act of 1936 [or the Anti-Price Discrimination Act, 15 U.S.C. § 13]) that preclude firm-specific pricing (price discrimination) for commodity inputs. Even if state-contingent pricing were feasible, it would still be cheaper for the firm to self-insure with cash. To understand this, assume that a factor requires a payment of K next period. If his or her contract were traded, the factor could agree to different payments in different states such that he or she can get a payment of E(K). As a security, this factor contract would be priced at zero NPV. However, because the factor contract is not traded, he or she will demand a higher expected return—an expected price K* next period, with $E(K^*) > E(K)$. The factor contract will then be a negative NPV investment to the firm. Thus, the firm is better off by self-insuring with cash (this will be a zero NPV investment).

REFERENCES

Ambler, Tim and John Roberts (2006), "Beware the Silver Metric: Marketing Performance Measurement Has to Be Multidimensional," Report No. 06-113, Marketing Science Institute.

Anderson, Eugene, Claes Fornell, and Sanal K. Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value," *Journal of Marketing,* 68 (October), 172–85.

———, ———, and Roland T. Rust (1997), "Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services," *Marketing Science*, 16 (2), 129–45.

Anderson, Paul F. (1982), "Marketing, Strategic Planning and the Theory of the Firm," *Journal of Marketing*, 46 (Spring), 15–26.

Arrow, Kenneth J. and Frank H. Hahn (1971), *General Competitive Analysis* San Francisco: Holden-Day Inc.

Brealey, Richard A., Stewart C. Myers, and Franklin Allen (2006), *Principles of Corporate Finance*, 8th ed. New York· McGraw-Hill/Irwin.

Coase, Ronald H (1992), "The Institutional Structure of Production," *The American Economic Review*, 82 (September), 713–19.

Cochrane, John H. (2001), *Asset Pricing*. Princeton, NJ: Princeton University Press.

Constantinides, George M. (2002), "Rational Asset Prices," *Journal of Finance*, 57 (August), 1567–92.

Copeland, Thomas E. and Fred Weston (1979), *Financial Theory and Corporate Policy*. Reading, MA: Addison-Wesley.

Day, George S. and Liam Fahey (1988), "Valuing Marketing Strategies," *Journal of Marketing*, 52 (July), 45–57.

Dekimpe, Marnik G. and Dominique M. Hanssens (1995), "The Persistence of Marketing Effects on Sales," *Marketing Science*, 14 (1), 1–21.

——— and ——— (1999), "Sustained Spending and Persistent Response: A New Look at Long-Term Marketing Profitability," *Journal of Marketing Research*, 36 (November), 397–412.

Doyle, Peter (2000), *Value-Based Marketing: Marketing Strategies for Corporate Growth and Shareholder Value*. Chichester, UK: John Wiley & Sons

Fama, Eugene F. and Merton H. Miller (1972), *The Theory of Finance*. Hinsdale, IL: Dryden Press.

Fornell, Claes, Sunil Mithas, Forrest V. Morgeson III, and M.S. Krishnan (2006), "Customer Satisfaction and Stock Prices," *Journal of Marketing*, 70 (January), 3–14.

Grossman, Sanford (1995), "Dynamic Asset Allocation and the Informational Efficiency of Markets," *Journal of Finance*, 50 (June), 773–87.

Gruca, Thomas S. and Lopo L. Rego (2005), "Customer Satisfaction, Cash Flow, and Shareholder Value," *Journal of Marketing*, 69 (July), 115–30.

Gupta, Sunil and Donald R. Lehmann (2005), *Managing Customers as Investments: The Strategic Value of Customers in the Long Run*. Upper Saddle River, NJ: Pearson Education.

———, ———, and Jennifer A. Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (February), 7–18.

——— and Valarie Zeithaml (2006), "Customer Metrics and Their Impact on Financial Performance," *Marketing Science*, 25 (6), 718–39.

Heins, A. James and Case A. Sprenkle (1969), "A Comment on the Modigliani-Miller Cost of Capital Thesis," *The American Economic Review*, 59 (4), 590–92.

Hunt, Shelby D. (2002), *Foundations of Marketing Theory*. Armonk, NY: M.E. Sharpe.

Kotler, Philip and Sidney Levy (1969), "A New Form of Marketing Myopia: Rejoinder to Professor Luck," *Journal of Marketing*, 33 (July), 55–57.

Kumar, Nirmalya (2004), *Marketing as Strategy*. Boston: Harvard Business School Press.

Lamey, Lien, Barbara Deleersnyder, Marnik G. Dekimpe, and Jan-Benedict E.M. Steenkamp (2007), "How Business Cycles Contribute to Private-Label Success: Evidence from the United States and Europe," *Journal of Marketing*, 71 (January), 1–15.

Lehmann, Donald R. (2004), "Linking Marketing to Financial Performance and Firm Value," *Journal of Marketing*, 68 (October), 73–75.

LeRoy, Stephen F. and Jan Werner (2001), *Principles of Financial Economics*. Cambridge, UK: Cambridge University Press.

Magill, Michael and Martine Quinzii (1996), *The Theory of Incomplete Markets*, Vol. 1. Cambridge, MA: The MIT Press.

McAlister, Leigh, Raji Srinivasan, and MinChung Kim (2007), "Advertising, Research and Development, and Systematic Risk of the Firm," *Journal of Marketing*, 71 (January), 35–48.

Modigliani, Franco and Merton H. Miller (1958), "The Cost of Capital, Corporation Finance, and the Theory of Investment," *The American Economic Review*, 48 (3), 261–97.

——— and ——— (1969), "Reply to Heins and Sprenkle," *The American Economic Review*, 59 (4), 592–95.

Moorman, Christine and Roland T. Rust (1999), "The Role of Marketing," *Journal of Marketing*, 63 (Special Issue), 180–97.

Nath, Pravin and Vijay Mahajan (2008), "Chief Marketing Officers: A Study of Their Presence in Firms' Top Management Teams," *Journal of Marketing*, 72 (January), 65–81.

Rappaport, Alfred (1986), *Creating Shareholder Value: The New Standard for Business Performance*. New York: The Free Press.

Rust, Roland T., Tim Ambler, Gregory S. Carpenter, V. Kumar, and Rajendra K. Srivastava (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (October), 76–89.

———, Katherine N. Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (January), 109–127.

———, Anthony Zahorik, and Timothy Keiningham (1995), "Return on Quality (ROQ): Making Service Quality Financially Accountable," *Journal of Marketing*, 59 (April), 58–70.

Shiller, Robert J. (1993), *Macro Markets*, Clarendon Lectures in Economics. Oxford: Oxford University Press.

——— (2003), *The New Financial Order*. Princeton, NJ: Princeton University Press.

Smithburg, William (1993), "Creating Value," interview with Brian McWilliams, *Enterprise*, (April), (accessed August 6, 2006), [available at http://www.snc.edu/socsci/chair/333/eva_2.htm].

Srivastava, Rajendra and David J. Reibstein (2005), "Metrics for Linking Marketing to Financial Performance," Report No. 05-200, Marketing Science Institute.

———, Tasadduq A. Shervani, and Liam Fahey (1997), "Driving Shareholder Value: The Role of Marketing in Reducing Vulnerability and Volatility of Cash Flows," *Journal of Market Focused Management*, 2 (1), 49–64.

———, ———, and ——— (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis," *Journal of Marketing*, 62 (January), 2–18.

Stewart, David W. (2006), "Making Marketing Accountable: The Role of Standards and Metrics," paper presented at the American Marketing Association Winter Educators' Conference, St. Petersburg, FL (February 18).

Webster, Frederick E., Jr., Alan J. Malter, and Shankar Ganesan (2005), "The Decline and Dispersion of Marketing Competence," *Sloan Management Review*, 46 (Summer), 35–43.

Girish Ramani & V. Kumar

Interaction Orientation and Firm Performance

Marketing managers are being required to demonstrate the profitability of their marketing actions down to the level of their individual customers and on an ongoing basis. At the same time, customers expect firms to increasingly customize their products and services to meet their demands. Firms still need to produce superior products, sell smarter, and understand the markets as a whole, but the ability of firms to orient themselves to interact successfully with their individual customers will differentiate them in the future. Advances in technology have resulted in increasing opportunities for interactions between firms and customers, between customers, and between firms. An interaction orientation reflects a firm's ability to interact with its individual customers and to take advantage of information obtained from them through successive interactions to achieve profitable customer relationships. First, the authors identify the components of interaction orientation: (1) customer concept, (2) interaction response capacity, (3) customer empowerment, and (4) customer value management. Second, they relate interaction orientation to both customer-level and aggregate-level performance measures. Third, they identify the antecedents of interaction orientation. Fourth, they examine the moderating effects of customer-initiated contacts and competitive intensity on the interaction orientation–performance linkage. The results are based on a survey of top marketing managers. The commonly held view that customer-based relational performance is related to customer-based profit performance is not supported. However, both customer-based relational performance and customer-based profit performance affect aggregate business-level performance positively. Interaction orientation is a phenomenon observed in both business-to-business and business-to-consumer firms. The extent of customer-initiated contacts moderates the interaction orientation–performance relationship.

Keywords: interaction orientation, structural equation modeling, customer value management, organization capabilities, firm performance

Marketing managers are being required to demonstrate the profitability of their marketing actions down to the level of their individual customers and on an ongoing basis. At the same time, customers expect firms to increasingly customize their products and services to meet their demands. Technological advances have heightened interactivity between customers and firms, customers and customers, and firms and firms (Yadav and Varadarajan 2005). Thus, increasing profit pressures, customer demand heterogeneity, and advances in technology all suggest that firms need to develop an orientation that is appropriate for survival and success in increasingly interactive market environments. Interactions help firms refine their knowledge about customer tastes and preferences (Srinivasan, Anderson, and Ponnavolu 2002). The effective and efficient management of interactions and the interfaces at which these interactions occur are increasingly being recognized as sources of lasting competitive advantage (Rayport and Jaworski 2005).

We believe that an interaction orientation reflects a firm's ability to interact with its individual customers and to take advantage of information obtained from them through successive interactions to achieve profitable customer relationships. However, no comprehensive construct exists in the literature that captures the key elements of an interaction orientation. We address this research gap. We develop interaction orientation as a composite construct that consists of an organization's fundamental belief (the customer concept) and the relevant processes (interaction response capacity) and practices (customer empowerment and customer value management) that supplement this fundamental belief. Thus, at the core of interaction orientation is the specific notion of the customer concept—the belief that the unit of analysis of every marketing action and reaction should be the individual customer.

We examine two composite customer-centric performance measures: (1) customer-based relational performance, which consists of customer satisfaction, customer

---

Girish Ramani is an assistant professor, LeBow College of Business, Marketing Department, Drexel University (e-mail: girish.ramani@drexel.edu). V. Kumar is ING Chair Professor and Executive Director, ING Center for Financial Services, School of Business Administration, Department of Marketing, University of Connecticut (e-mail: Vk@business.uconn.edu). The authors thank the three anonymous *JM* reviewers for their valuable comments on a previous version of this article. They thank the Marketing Science Institute for providing access to the marketing executives in various companies to obtain the necessary data for conducting this research study. The authors also benefited from interacting with the executives in the American Marketing Association, the Research Board, and the Conference Board conferences in obtaining the data for the study. They also thank Ajay Kohli, John Hulland, Stan Slater, and Christine Moorman as well as the participants at the 2006 American Marketing Association Winter Educators' Conference and the 2006 Marketing Science Conference. Finally, they thank the Marketing Science Institute for supporting this study with financial assistance.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

ownership, and positive word of mouth, and (2) customer-based profit performance, which consists of the successful identification of profitable customers, efficiency of the acquisition and retention process, and conversion of unprofitable customers to profitable ones. Our hypotheses posit a positive primary impact of interaction orientation on the two customer-based performance measures and a positive relationship between customer-based relational performance and customer-based profit performance. We also hypothesize five antecedents of interaction orientation: the extent of dependence on patents and trademarks for success, the extent of institutional pressures to adopt interactive technologies, the extent to which an employee reward system is based on customer-centric performance measures, the degree of outsourcing expertise possessed by the firm, and whether the firm operates in the business-to-business industry. Finally, we hypothesize moderating effects of customer-initiated contacts and competitive intensity on the two interaction orientation–customer-based performance links.

We develop new scale items and adapt existing scales to ascertain a firm's level of interaction orientation and customer- and aggregate-level performance, as well as where the firm stands on the hypothesized antecedent factors. We describe the survey-based data collection process and analyze the data using confirmatory factor analysis (CFA), structural equation modeling, and moderator regression analyses. We compare the hypothesized model with a rival model. We also analyze the effects of customer-based relational performance and customer-based profit performance on aggregate business-level performance. We offer managerial implications for firms that adopt and do not adopt an interaction orientation. Our study establishes that an interaction orientation generates value for a firm, is rare, and therefore is a marketing resource that leads to a position of competitive advantage (cf. Hunt and Morgan 1995).

Interaction Orientation: Development of the Construct

We conducted exploratory interviews with 48 managers from 26 business-to-business and 18 business-to-consumer firms. There were 2 respondents from 4 firms and 1 respondent each from the remaining 40 firms. The choice of both types of firms is in line with contemporary marketing thought that interactive marketing is now relevant to all types of organizations (Coviello et al. 2002; Day and Montgomery 1999; Grönroos 1994).

We now summarize the responses to the two key questions asked during the interviews: (1) Is marketing more about interactions with individual customers than before? and (2) What are the capabilities that firms require to offer superior interactions? Most managers voiced the need to move away from a market segment to an individual customer approach when analyzing the effect of marketing actions. We term this belief the "customer concept." Some managers believed that managing individual customers entails using sophisticated database systems. Others believed that customer-facing employees need to adapt to individual customer needs. This suggested that the belief in the customer concept must be supplemented with appropriate systems. Thus, we termed the processes and systems the firm adopts to interact with and respond to its customers "interaction response capacity." Managers also believed that individual customers should be allowed a greater say in setting the terms of interactions with the firm and with each other. Thus, a firm should demonstrate that customers are important partners when they interact with the firm and even when they interact with each other. We term this practice "customer empowerment." Managers also believed that there is a growing understanding among customers that the firm has the right to treat individual customers differently according to their value to the firm. Therefore, firms should develop the ability to break down the revenues and costs related to each firm–customer interaction. We term this practice "customer value management."

Thus, on the basis of the feedback from our interviews, we propose interaction orientation as a composite construct that captures (1) a firm's belief in the customer concept, (2) a firm's interaction response capacity that reflects its ability to use dynamic database systems and processes, (3) a firm's customer empowerment practices that help shape customer–firm interactions and customer–customer interactions, and (4) a firm's customer value management practices that guide its marketing resource allocation decisions. Whereas most conceptualizations of a market orientation limit themselves to either a behavioral or a cultural perspective, we take the view that a comprehensive construct of interaction orientation should capture the basic underlying belief and also the processes and practices that supplement this belief (see Homburg and Pflesser 2000). We now draw from extant literature and offer precise definitions for the four elements of interaction orientation.

The Customer Concept

The customer concept is characterized by the individual customer as the starting point for marketing activities (Hoekstra, Leeflang, and Wittink 1999). We define the customer concept as the belief that prescribes the unit of analysis of every marketing action and reaction to be the individual customer.

Interaction Response Capacity

Interaction response capacity represents the degree to which the firm offers successive products, services, and relationship experiences to each customer by dynamically incorporating feedback from previous behavioral responses of that specific customer and of other customers collectively. Thus, it reflects the ability of a firm's systems to respond to heterogeneous customers differently and also to each individual customer differently at different points in time by pooling information from multiple sources and points in time.

Customer Empowerment

A theory gaining credence is that firms cannot think and act unilaterally, and in this new paradigm of marketing, the consumer and the firm cocreate value at various points of interaction (Prahalad and Ramaswamy 2004). Customer empowerment reflects the extent to which a firm provides its customers avenues to (1) connect with the firm and actively shape the nature of transactions and (2) connect and collaborate with each other by sharing information; praise; criticism; suggestions; and ideas about its products, services, and policies.

Customer Value Management

To reap the economic benefits of available customer data, firms need to develop the practice of aligning resources spent on customers in proportion to the revenues or profits derived from them (Mulhern 1999; Reinartz, Krafft, and Hoyer 2004). Data-analytic techniques that enable the measurement and prediction of customer-based revenues and profits (e.g., the recency, frequency, and monetary value; past customer value; customer lifetime value) are already emerging as key topics in marketing textbooks (e.g., Kumar and Reinartz 2006). Relationship management based on customer value involves providing differentially tailored treatment according to the expected response from each customer to available marketing initiatives, such that the contribution from each customer to overall profitability is maximized (Kumar, Ramani, and Bohling 2004). In other words, customer value management represents the extent to which the firm can define and dynamically measure individual customer value and use it as its guiding metric for marketing resource allocation decisions.

We now examine interaction orientation with respect to related concepts available in the literature, such as relationship orientation and customer-relating capability. The term "relationship orientation" has been conceptualized as being the opposite of a transaction mentality (Day 1999, 2000); it "reflects relevant values, behavioral norms, the shared mental modes used to make sense out of patterns of customer loyalty and defection, and decision criteria" (Day and Van den Bulte 2002, pp. 7–8). In contrast, interaction orientation is precisely defined in terms of its four specific components. Interaction orientation is also different from the broader concept of customer-relating capability (see Day and Van den Bulte 2002) in that interaction orientation is specific and actionable and can be adopted by firms to achieve superior performance.

The Consequences of Interaction Orientation

The consequences of firm-level strategic orientations, such as market orientation, selling orientation, and production orientation, have been typically evaluated in terms of aggregate business-level performance measures, such as return on sales and return on assets (Noble, Sinha, and Kumar 2002), as well as sales, profits, and market shares (Voss and Voss 2000). However, customer-centric organizations have different methodologies, vocabulary, metrics, and evaluation criteria than product-centric organizations (Sheth, Sisodia, and Sharma 2000). Researchers have recently begun to examine the impact of marketing efforts in terms of customer-centric measures (e.g., Jayachandran et al. 2005). Because interaction orientation is based on the belief that the individual customer should be the unit of analysis, it would be logical to examine whether firms that are interaction oriented indeed exhibit superior performance in terms of customer-centric measures. We examine the consequences of an interaction orientation on two groups of customer-centric performance measures: (1) customer-based relational performance and (2) customer-based profit performance (see Figure 1). Customer-based relational performance assesses performance on attitudinal parameters, whereas customer-based profit performance assesses performance on behavioral parameters.

Customer-Based Relational Performance

We specify a measure of customer-based relational performance in terms of three indicators: (1) customer satisfaction, (2) customer ownership, and (3) positive word of mouth. Satisfaction indexes provide an indication of the strength of the relationship between a firm and its customers. A superior interaction response capacity and consistent customer empowerment practices are likely to result in greater customer satisfaction. Customer ownership refers to the degree to which customers feel accountable to a firm and actively seek the firm's financial well-being. Firms that resist the growth of power in the hands of their customers risk distancing themselves from their source of business. A successful business strategy is likely to be one that empowers individual customers by allowing them to develop experiences with the company on their terms (Newell 2003; Prahalad and Ramaswamy 2004). Customers who are empowered and rewarded according to their individual expertise and needs develop a greater sense of belonging to the firm and are likely to protect the well-being of the firm. In other words, conscious efforts by a firm to develop and enhance an interaction orientation will result in greater customer ownership of the firm. Word of mouth refers to the spread of information about products, services, stores, companies, sales, or customer managers from one customer to another (Brown et al. 2005). Customers who make a personal referral must not only believe that a company offers superior economic value but also feel good about their relationship with the company (Reichheld 2006). An interaction orientation increases positive word of mouth by encouraging and enabling customers to refer the firm to new customers and new customers to the firm.

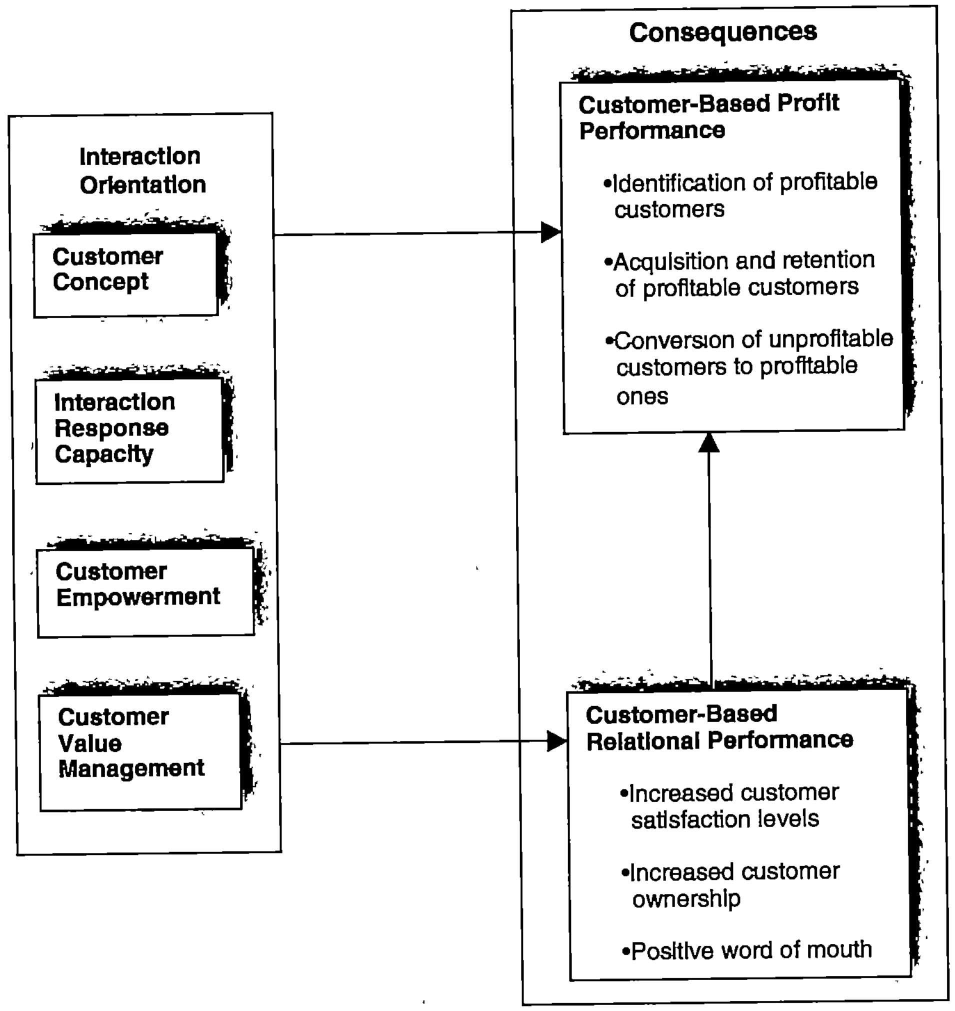
Thus, we propose that an interaction orientation leads to superior customer-based relational performance, which is measured in terms of customer satisfaction, customer ownership, and positive word of mouth. Formally,

> H_1: The greater the interaction orientation of a firm, the greater is its customer-based relational performance.

Customer-Based Profit Performance

We specify a measure of customer-based profit performance in terms of three indicators: (1) identification of profitable customers, (2) acquisition and retention of profitable customers, and (3) conversion of unprofitable customers to profitable ones. An interaction orientation facilitates the analysis of transaction data obtained from various interfaces and, consequently, the use of this information to relate individual customer revenues to marketing investments. Measuring customer-level profits ensures marketing managers'

FIGURE 1
The Customer-Level Consequences of Interaction Orientation


Interaction Orientation
Customer Concept
Interaction Response Capacity
Customer Empowerment
Customer Value Management
Consequences
Customer-Based Profit Performance
•Identification of profitable customers
•Acquisition and retention of profitable customers
•Conversion of unprofitable customers to profitable ones
Customer-Based Relational Performance
•Increased customer satisfaction levels
•Increased customer ownership
•Positive word of mouth

accountability while increasing a firm's agility to provide effective and efficient marketing responses to customer signals. Firms that invest in processes that enhance their interaction response capacity and implement customer value management practices are in a position to distinguish between the characteristics of profitable customers and those of unprofitable customers and to use this information to identify potentially profitable customers (e.g., Reinartz and Kumar 2003). Firms that adopt customer value management principles also realize that focusing on the long-term value of customers, and not simply maximizing either acquisition or retention, leads to superior overall efficiencies (Thomas, Reinartz, and Kumar 2004).

A customer becomes unprofitable when the costs of acquiring, selling, and retaining that customer exceeds the revenue contribution from the customer over time. Using modeling techniques, firms that identify profitable customers could also determine variables that drive customer profitability (e.g., Reinartz and Kumar 2003). By dynamically capturing individual customer profitability, it is possible to monitor cost and revenue variables that are under the control of the firm. Studies have shown that the customer lifetime value metric helps firms plan suitable marketing and communication channel mixes and provide time- and product-based cross-selling and up-selling recommendations for individual customers (Kumar, Ramani, and Bohling 2004; Venkatesan and Kumar 2004).

We synthesize the preceding arguments to propose that an interaction orientation leads to superior customer-based profit performance, measured in terms of the degree of success in identifying profitable customers, balancing the acquisition and retention of customers profitably, and converting unprofitable customers to profitable ones. Thus:

H_2: The greater the interaction orientation of a firm, the greater is its customer-based profit performance.

Customer-Based Relational Performance and Customer-Based Profit Performance

Satisfaction scores have been linked to increased business from customers (Anderson, Fornell, and Mazvancheryl 2004). Negative word-of mouth behaviors could have detrimental effects on the value of a firm's customer base (Hogan, Lemon, and Libai 2003), whereas customers indulging in positive word of mouth are of great importance to the firm (Kumar, Petersen, and Leone 2006; Reichheld 2006). In other words, measures of customer-based relational performance, such as satisfaction and word of mouth, are positively related to measures of customer-based profitability. Thus, on the basis of the generally accepted notion that superior relational performance leads to superior profit performance, we offer the following hypothesis:

H_3: The greater a firm's customer-based relational performance, the greater is its customer-based profit performance.

The Antecedents of Interaction Orientation

The notion of interaction orientation resonates with the managers contacted in our exploratory interviews, but these managers also believed that several factors could affect the level of interaction orientation a firm exhibits—for example, the outlook and prior experience of its top management, the nature of the firm's existing business, and the prevalent industry and competitive practices. Therefore, we classify our antecedents in terms of management-level, firm-level, and industry-level characteristics. The employee reward system is a management-level factor that has been examined for its effect on market orientation (Jaworski and Kohli 1993) and on marketing strategy comprehensiveness (Atuahene-Gima and Murray 2004), and it is also likely to affect the firm's interaction orientation. Firm characteristics and industry characteristics have been examined in contingency models for their effect on competitive strategy (Varadarajan and Yadav 2002) and competitive positional advantage (Bharadwaj, Varadarajan, and Fahy 1993). In increasingly interactive market environments, dependence on patents and expertise in outsourcing are firm-level characteristics that influence the firm's interaction orientation, and institutional pressures and industry type are industry-level characteristics that could influence the firm's degree of interaction orientation. We now develop hypotheses that relate these antecedents to interaction orientation.

Jaworski and Kohli (1993) show that a market-oriented reward system has a strong impact on market orientation. An interaction-oriented reward system focuses on customer-level performance measures. If a firm's employee reward system is based on customer metrics (e.g., customer acquisition, customer retention, customer win-back, customer profitability) instead of sales and market share measures, we could expect a greater degree of adoption of customer value management practices. Thus:

H_4: The greater a firm's reliance on customer metrics for evaluating and rewarding managers, the greater is its interaction orientation.

Recently, IBM provided free access to 500 patents to companies, groups, or individuals working on open-source projects in a move welcomed by academics and industry analysts (Lohr 2005). The use of patents to enjoy a short-term lead is fading in importance in many industries, except for the pharmaceutical industry (Reitzig 2004). Firms that are ensured business by virtue of the patents they own on their products are less likely to invest in processes to interact with their customers. Dependence on patents fosters a product-centric management approach in firms. If firms depend on patents to sustain their business, they might not recognize customers' needs to bond with their products and are unlikely to develop a high degree of interaction orientation. Such firms resist the move to involve their customers because, in their current wisdom, they believe that there are limited returns to such a strategy. Thus:

H_5: The lower the dependence on trademark and patent protection as the source of continued business for a firm, the greater is its interaction orientation.

Rather than settling for the subset of prospects who find one offer relevant, it is better for firms to attract and retain customers by presenting many relevant offerings to each customer (Newell 2003). Outsourcing vastly expands the ability to provide a wide range of products and services, but success in outsourcing depends on the degree to which contracts with suppliers and the performance levels expected of suppliers are specified and monitored (King 2004). A superior control of outsourced back-end supply systems enhances a firm's interaction response capacity. Thus:

H_6: The greater the outsourcing expertise of a firm, the greater is its interaction orientation.

Organizations that embrace electronic markets to mimic a successful benchmark firm believe that the benchmarked organization succeeded primarily because of its participation in electronic markets (Grewal, Comer, and Mehta 2001). Normative pressure caused by the sheer number of competitive firms adopting new interactive technologies hastens a firm's adoption of interactive tools (Tsikriktsis, Lanzolla, and Frohlich 2004; Wu, Mahajan, and Balasubramanian 2003). Thus:

H_7: The greater the institutional pressures for a firm to adopt interactive technologies, the greater is its interaction orientation.

Business-to-business and business-to-retail firms organize themselves into account management teams that service individual clients. Therefore, we expect a greater acceptance and dissemination of the belief in the customer concept and the adoption of processes and practices commensurate with this belief in business-to-business firms than in business-to-consumer firms. Thus:

H_8: Business-to-business firms exhibit a greater degree of interaction orientation than business-to-consumer firms.

Moderators: Customer-Initiated Contacts and Competitive Intensity

Given that reciprocity of firm–customer communication is integral to interaction orientation, we needed a moderator variable that could capture the differences between firms in terms of the behaviors of their customer bases toward the firm. Interactions drive relationships, but interactions do not constitute a genuine relationship unless the customer acknowledges that they do (Peppers and Rogers 2004). The extent of two-way communication determines the strength of the firm–customer relationship (Venkatesan and Kumar 2004). A stringent measure of customers' acknowledgment of a firm's attempt to forge a genuine relationship would be the extent to which a firm's customers initiate communications with the firm. We identified customer-initiated contacts as a moderator variable on the basis of the importance established for it in the marketing literature in the business-to-business area. Customer-initiated contacts provide rich information about a customer's needs and concerns to a firm and enable the firm to interact with the customer and tailor appropriate responses without incurring heavy expenses (Bowman and Narayandas 2001). Moreover, the responsiveness to factors under a firm's control varies across customer-initiated contacts, and therefore firms that adapt their processing have an advantage (Bowman and Narayandas 2001). A firm that has a customer base that exhibits low levels of customer-initiated contacts is not likely to experience significantly improved customer-based profit performance outcomes, despite a high level of interaction orientation. Thus:

> H_9: The higher the level of customer-initiated contacts, the greater is the positive effect of a firm's interaction orientation on its customer-based profit performance.

We chose competitive intensity as the second moderator variable because it is the most frequently studied moderator in research that links market orientation and performance (Kirca, Jayachandran, and Bearden 2005). In situations of high competitive intensity, competitors erode a firm's product-based advantage by imitating or improving the product offerings. Firms that possess a high level of interaction orientation are likely to differentiate themselves from their competition not in terms of products but rather by focusing on individual customers' characteristics and needs. As competitive intensity increases, the effect of interaction orientation on customer satisfaction, customer ownership, and positive word of mouth should increase. Thus:

> H_{10}: The greater the competitive intensity, the greater is the positive effect of a firm's interaction orientation on its customer-based relational performance.

Research Methodology

We developed a structured survey instrument in several stages. We designed the scale items to measure practicing managers' perceptions of their respective firms.

Measure Development

We developed new scales for the four dimensions of the interaction orientation (INTOR) construct, as well as the customer-based relational performance (CBRP), customer-based profit performance (CBPP), and the four antecedent constructs. We also included measures for aggregate business-level performance (BLP), competitive intensity (COMPIN), and customer-initiated contacts (CICs) after suitably adapting existing scales.

Following the framework proposed by Churchill (1979), we generated an item pool for each construct. We used literature in market orientation (e.g., Kohli, Jaworski, and Kumar 1993) and customer relationship management (CRM) (e.g., Jayachandran et al. 2005; Reinartz, Krafft, and Hoyer 2004) as guidance to develop the item pool and for refining the scales. We then pretested the scale items with six senior marketing executives for comprehension, logic, and relevance. On the basis of the feedback, we modified the scale items and presented them to a panel of four academic experts in CRM to examine face validity. A questionnaire containing 46 items resulted. We administered the 46-item questionnaire to a test pool of 51 senior marketing executives. We conducted an exploratory factor analysis of their responses. We examined coefficient alphas and used exploratory factor analyses to purify our scales. We discarded 12 items because of the item-to-total correlations and the factor loadings, using a cutoff value of .4 for both, which left 34 indicators. The number of indicators we used in this study is similar to the number of indicators used in other studies that examine latent constructs derived from surveys conducted with key informants in firms (e.g., Jayachandran et al. 2005). Eight questions elicited information on business type, industry category, COMPIN, CICs, annual sales, and employee size for the respondent firm.

As described previously, we conceptualized INTOR as a second-order construct consisting of four dimensions: customer concept (CC), interaction response capacity (IRC), customer empowerment (CE), and customer value management (CVM). The three CC items reflect the strength of the belief of a firm in viewing, acquiring, and analyzing customers at the individual level. The four IRC items measure the sophistication of a firm's systems to record, identify, access, and predict individual-level transactions. The three CE items record the extent to which a firm facilitates its customers to share feedback on its products and services with the firm and with other customers and to participate actively in designing products and services. The CVM scale consists of three items that measure the extent to which the firm has skills and processes in place to compute individual-level profits, predict individual-level profits, and record revenues accruing from individual customers for every marketing action.

We measured CBPP using three items that represented a firm's efficiency in predicting customer-level profitability, balancing acquisition and retention profitably, and converting unprofitable customers to profitable ones. We measured CBRP using three items that gauged the extent of customer satisfaction, customer ownership, and customer referrals in a firm.

We adapted three single-item measures of aggregate BLP from prior studies to obtain self-report measures on current profits relative to competitors' (BLP1), current profits relative to previous year (BLP2), and return on marketing investments (BLP3).

We developed two-item scales each for dependence on patents (PAT) and outsourcing expertise (OUTS). We adapted an existing scale (Wu, Mahajan, and Balasubramanian 2003) to provide a two-item measure for obtaining the extent of normative institutional pressure to adopt interactive technologies (INST). We used an existing scale (Jaworski and Kohli 1993) as a basis to develop a two-item scale to measure employee rewards systems (EMP) (see the Appendix).

Sample Characteristics and Data Collection

The respondents in this study were senior- and top-level marketing executives. We did not constrain the sample to specific industries in the interest of generalizability of our findings. We identified potential respondents from the membership lists of the Research Board, Conference Board, Marketing Science Institute, and the American Marketing Association. We followed a convenience sample approach to ensure that key informants were chosen as respondents on the basis of their knowledge of marketing activities and seniority in their respective organizations. Members of the Marketing Science Institute were sent e-mails directly by the institute with a request to participate in the survey. The contact lists were verified, and care was taken not to send duplicate requests. A total of 375 potential respondents representing 175 firms were contacted by e-mail with a request to participate in an online survey or to complete and return the attached electronic document that contained the survey. Two respondents were targeted in each of 120 firms, three respondents were targeted in each of 40 firms, and one respondent was targeted in each of 15 firms. Respondents were assured of the confidentiality of both respondent and firm identities. Second and third reminder e-mails were sent one week and two weeks, respectively, after the first one.

We received responses from 125 firms. The average time taken to view the online questionnaire was 11 minutes. We used 211 complete responses, representing 107 firms, in the analysis, after we performed a listwise deletion procedure. At the firm level, this represented a 61% response rate. Of the firms, 74 were from business-to-business industries, and the remaining 33 were from business-to-consumer industries. The annual sales of these firms ranged from $100 million to $90 billion. The average sales value was $5 billion. A comparison of the averages of annual sales and the number of employees between early and late respondents indicated no statistically significant differences at the .05 level.

An examination of interrater correlations on the scale items, when we had at least two respondents from the same firm, indicated that they ranged from .8 to .9. This reflects minor differences among informant reports.[1] Thus, we used the unweighted average of the responses to arrive at composite measures (see Van Bruggen, Lilien, and Kacker 2002; Kumar, Stern, and Anderson 1993). In other words, when there were multiple respondents from the same firm, we computed the average of scale items obtained from multiple respondents and used the average to represent the firm. Thus, our unit of analysis in this study was the firm.

Analysis and Results

We examined all scale items and reverse-coded data when applicable to reflect the hypothesized directions. We used a combination of SPSS 13, AMOS 5.1, and Microsoft Excel 2003 software packages to carry out the analyses. We examined the univariate skewness and kurtosis of the variables and found them to be within acceptable levels.

Measurement Models

We evaluated measurement properties by running CFAs, first on the focal construct of INTOR and subsequently on the seven-factor model that included the two customer-based performance factors (CBPP and CBRP) and the four antecedent factors (PAT, OUTS, INST, and EMP). We used raw data as the input to this maximum likelihood–based estimation procedure. This approach is consistent with analyses dealing with organizational- and customer-level latent constructs (e.g., Kandemir, Yaprak, and Cavusgil 2006; Jayachandran et al. 2005; Jones and Reynolds 2006) and follows Gerbing and Anderson's (1988) guidelines.

The INTOR construct. As we discussed previously, INTOR is a second-order construct, and its four dimensions (CC, IRC, CE, and CVM) are first-order factors measured through their respective indicators. We conducted a CFA of this hierarchical model (see Figure 2).

The second-order CFA model fit was deemed to be acceptable on the basis of a battery of fit indexes (χ^2 = 167.14, d.f. = 61; comparative fit index [CFI] = .94; goodness-of-fit index [GFI] = .82; Tucker–Lewis index [TLI] = .92; incremental fit index [IFI] = .94; standardized root mean square residual [SRMR] = .05). The set of fit indexes reported is consistent with Hu and Bentler's (1999) recommendations. The path coefficients between the indicators and their respective first-order factors were significant at the α = .05 level. In addition, all the path coefficients between the second-order construct (INTOR) and its four dimensions (CC, IRC, CE, and CVM) were significant at the α = .05 level (see Table 1, Panel A).

We also examined the second-order factor structure by conducting a one-factor CFA on the average scores of the respective four first-order constructs (e.g., Jayachandran et al. 2005). The model fit was good (χ^2 = 14.59, d.f. = 2; CFI = .98; GFI = .95; TLI = .93; IFI = .98; and SRMR = .03). All the path coefficients were significant at the α = .05 level (see Table 1, Panel B). Thus, consistent with common practice (e.g., Jayachandran et al. 2005; Matsuno, Mentzer, and Ozsomer 2002), we used the aggregated scale consisting of the average scores of the four dimensions of INTOR as indicators of INTOR for further analyses.

The combined measurement model. Before estimating the path coefficients of the hypothesized structural model, we proceeded to fit a CFA on all the seven latent factors: INTOR, CBRP, CBPP, EMP, OUTS, INST, and PAT.[2]

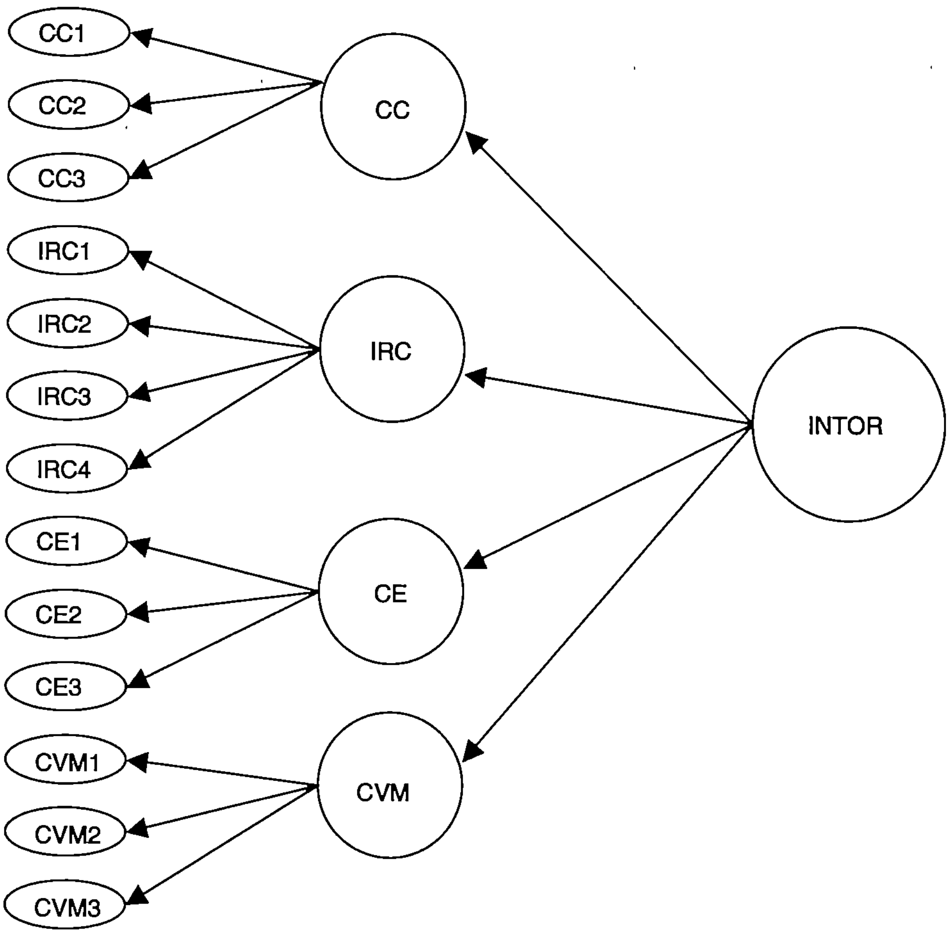
Composite reliability is an indicator of the shared variance among the set of observed variables used as indicators of a latent construct (Fornell and Larcker 1981; Kandemir, Yaprak, and Cavusgil 2006). As Table 2, Panel A, shows, construct reliabilities for all the seven latent constructs

---

[1] In response to a suggestion from an anonymous reviewer, we also examined the interrater agreement for each construct by computing the average deviation index. All computed values of the index were significantly below the cutoff value when tested using the software program recommended by Burke and Dunlap (2002) (available at www.tulane.edu/~dunlap/psylib.html). This suggests acceptable interrater agreement across all the constructs.

[2] At this stage of the analysis, we do not include the binary variable B2B because it does not represent a hypothesized latent construct.

FIGURE 2
INTOR: A Second-Order CFA



Notes: The respective indicators of CC, IRC, CE, and CVM are numbered serially (e.g., CC1, CC2, . . , CVM3).

ranged from .82 to .97, well above the recommended value. In addition, the coefficient alpha values were well above the threshold value of .7 that Nunnally (1978) recommends.

The seven-factor CFA model exhibited a good fit with the data (χ^2 = 253.95, d.f. = 131; CFI = .95; GFI = .82; TLI = .93; IFI = .95; and SRMR = .07). The standardized factor loadings ranged from .62 to greater than .90 and were statistically significant at the α = .95 level (see Table 2, Panel B). This provided the necessary evidence that all the constructs exhibited convergent validity.

We examined discriminant validity using a procedure suggested by Fornell and Larcker (1981) and widely followed by other researchers (e.g., De Wulf, Odekerken-Schroder, and Iacobucci 2001; Kandemir, Yaprak, and Cavusgil 2006). We computed the average variance extracted by the indicators corresponding to each of the seven factors and compared it with the highest variance that each factor shared with the other factors in the model. The average variance extracted for each factor was always greater than the highest shared variance (see Table 2, Panel A). Although Fornell and Larcker's procedure is considered a demanding test for discriminant validity (De Wulf, Odekerken-Schroder, and Iacobucci 2001; Grewal, Cote, and Baumgartner 2004), we also examined discriminant validity using an alternative procedure that Anderson and Gerbing (1988) recommend.[3] The chi-square values for the unconstrained models, which allowed each pair of constructs to covary freely, were always significantly lower than those of the constrained models, which constrained the estimated correlation for each pair of estimated constructs to one. (For example, for the pair of constructs OUTS and INTOR, whose estimated correlation is .88, the unconstrained model had a chi-square of 36.8 [d.f. = 8], and the constrained model had a chi-square of 124.6 [d.f. = 9]. The chi-square difference is significant at $p < .001$.) Similar results for the remaining pairs of constructs indicated that discriminant validity had indeed been achieved. In addition, because Fornell and Larcker's criterion is satisfied in our study, an inference error due to multicollinearity is also unlikely (Grewal, Cote, and Baumgartner 2004).

We used the Harmon's one-factor test to assess whether a single latent factor would account for all the manifest variables. This would indicate whether common method variance posed a serious threat to the interpretation of the findings from this study (Jayachandran and Varadarajan 2006; Kandemir, Yaprak, and Cavusgil 2006). The single-

---

[3] We thank an anonymous reviewer for suggesting this test.

TABLE 1
Results of the CFA

A: Using a Second-Order Conceptualization of INTOR

Indicator	Direction	Construct	Standardized Loading	Unstandardized Loading	SE	t-Value	p
CC1	←	CC	.90	1.00			
CC2	←	CC	.85	.76	.06	12.23	.00
CC3	←	CC	.86	.90	.07	12.66	.00
IRC1	←	IRC	.89	1.00			
IRC2	←	IRC	.91	1.01	.07	14.62	.00
IRC3	←	IRC	.87	.80	.06	13.14	.00
IRC4	←	IRC	.90	1.06	.07	14.48	.00
CE1	←	CE	.94	1.00			
CE2	←	CE	.88	.91	.06	15.52	.00
CE3	←	CE	.86	.75	.05	14.32	.00
CVM1	←	CVM	.90	1.00			
CVM2	←	CVM	.97	1.21	.07	18.46	.00
CVM3	←	CVM	.92	1.15	.07	15.67	.00
CC[a]	←	INTOR[b]	.93	1.00			
IRC[a]	←	INTOR[b]	.97	1.39	.11	12.59	.00
CE[a]	←	INTOR[b]	.96	1.16	.10	11.50	.00
CVM[a]	←	INTOR[b]	.99	1.07	.09	11.84	.00

B: Using Average Scores for the Four Dimensions of INTOR

Indicator	Direction	Construct	Standardized Loading	Unstandardized Loading	SE	t-Value	p
CC	←	INTOR	.88	1.00			
IRC	←	INTOR	.96	1.18	.07	16.57	.00
CE	←	INTOR	.92	1.38	.09	14.80	.00
CVM	←	INTOR	.94	1.48	.10	15.39	.00

[a] Second-order indicators.
[b] Second-order factor.
Notes: The respective indicators of CC, IRC, CE, and CVM are numbered serially (e.g., CC1, CC2, ..., CVM3).

factor model yielded a chi-square of 724.83 (d.f. = 153). We conducted a chi-square difference test against the hypothesized seven-factor model to assess the impact of common method variance. A significant difference between the chi-square values of the two models indicated that the fit in the one-dimensional model was significantly worse ($\Delta\chi^2 = 470.8$, Δd.f. = 22, $p < .01$) than it was in the measurement model. This provided preliminary evidence that the measurement model was robust to common method variance. In addition, we used the procedure that Lindell and Whitney (2001) recommend and Jayachandran and colleagues (2005) adopt to test comprehensively for common method bias. We chose COMPIN, a moderator variable in our study, as the marker variable for the common method bias analysis because it is theoretically unrelated to the dependent variable CBRP. In our study, COMPIN and CBRP had a nonsignificant correlation of .17. Therefore, we used COMPIN's measured correlation with the dependent variable CBRP as the indication of method variance. The results, along with the reported correlations between all the constructs in the measurement model and COMPIN, appear in Table 3.

Table 3 shows that the partial correlations between the constructs hypothesized to have a significant relationship are significant even after we partial out the effect of common method bias. We also carried out a 95% sensitivity analysis to validate this result further. We concluded that the measurement model possessed good overall fit with the data, that the constructs exhibited both convergent and divergent validity, and that common method bias did not pose a serious threat to the interpretation of the results from this study.

The Structural Model

The structural model used to test the hypotheses consisted of all the seven factors tested in the measurement model and a single-item binary variable (B2B) that indicated the nature of the business of the respondent firm (see Figure 3). Categorical cause indicators, such as the binary variable B2B, have been used in multiple-indicators-and-multiple-causes structural models to estimate group differences on latent variables (Kline 2005). The model fit measures indicated acceptable agreement with the covariance in the data ($\chi^2 = 302.10$, d.f. = 153; CFI = .94; GFI = .81; TLI = .92; IFI = .94; and SRMR = .07).[4] The results of the hypothesis testing for H_1–H_8 appear in Table 4, Panel A.

We found that INTOR was positively associated with CBRP ($\beta = 1.19$, $p < .01$) and CBPP ($\beta = .77$, $p < .01$), in support of H_1 and H_2. H_3, which predicted that superior CBRP would lead to superior CBPP, was not supported ($\beta =$

---

[4] The model fit did not differ significantly when the size of the firm was included as a control variable in the analysis. Thus, we did not include the size variable in our model.

TABLE 2
Measurement Model

A: Scale Properties of the Seven Latent Factors

Construct	Construct Reliability	Coefficient Alpha	Average Variance Extracted	Highest Shared Variance
INTOR	.96	.96	.86	.78
CBPP	.93	.92	.81	.74
CBRP	.89	.88	.74	.70
PAT	.89	.88	.80	.57
EMP	.88	.86	.71	.55
INST	.82	.82	.69	.64
OUTS	.97	.97	.94	.78

B: Results of the CFA with the Seven Latent Factors

Indicator	Direction	Construct	Estimate	Standardized Estimate	SE	t-Value	*p*
CC	←	INTOR	1.00	.90			
IRC	←	INTOR	1.13	.94	.07	16.61	.00
CE	←	INTOR	1.37	.94	.08	16.31	.00
CVM	←	INTOR	1.44	.94	.09	16.25	.00
CBPP1	←	CBPP	1.00	.92			
CBPP2	←	CBPP	.74	.86	.06	13.55	.00
CBPP3	←	CBPP	1.03	.92	.07	15.64	.00
CBRP3	←	CBRP	1.00	.97			
CBRP1	←	CBRP	.58	.75	.06	10.24	.00
CBRP2	←	CBRP	.86	.85	.06	13.92	.00
EMP1	←	EMP	1.00	.62			
EMP2	←	EMP	1.74	.94	.23	7.48	.00
EMP3	←	EMP	1.73	.93	.23	7.39	.00
INST1	←	INST	1.00	.81			
INST2	←	INST	1.09	.85	.12	8.86	.00
PAT1	←	PAT	1.00	.97			
PAT2	←	PAT	.94	.82	.09	10.28	.00
OUTS2	←	OUTS	1.00	.99			
OUTS1	←	OUTS	.99	.95	.04	25.45	.00

.09, $p > .05$). In support of H_4, EMP was positively associated with INTOR ($\beta = .31$, $p < .01$), and in support of H_5, PAT was negatively associated with INTOR ($\beta = -.22$, $p < .01$). In support of H_6, OUTS was positively linked to INTOR ($\beta = .31$, $p < .01$), and in support of H_7, INST was positively linked to INTOR ($\beta = .25$, $p < .05$). H_8, which predicted a higher INTOR for business-to-business firms, was not supported ($\beta = .07$, $p > .05$). Overall, six of the eight proposed hypotheses tested through the structural model received support.

To evaluate the validity of the findings further, we carried out an analysis comparing the model fit of 10 randomly chosen subsamples, consisting of 90 respondents each, from our total sample of 107. Nonsignificant differences in the chi-square value when we placed equality constraints for the parameters across two subsamples at a time indicated no significant change in the model fit. This suggests that the model was valid across different subsamples.

The Rival Model

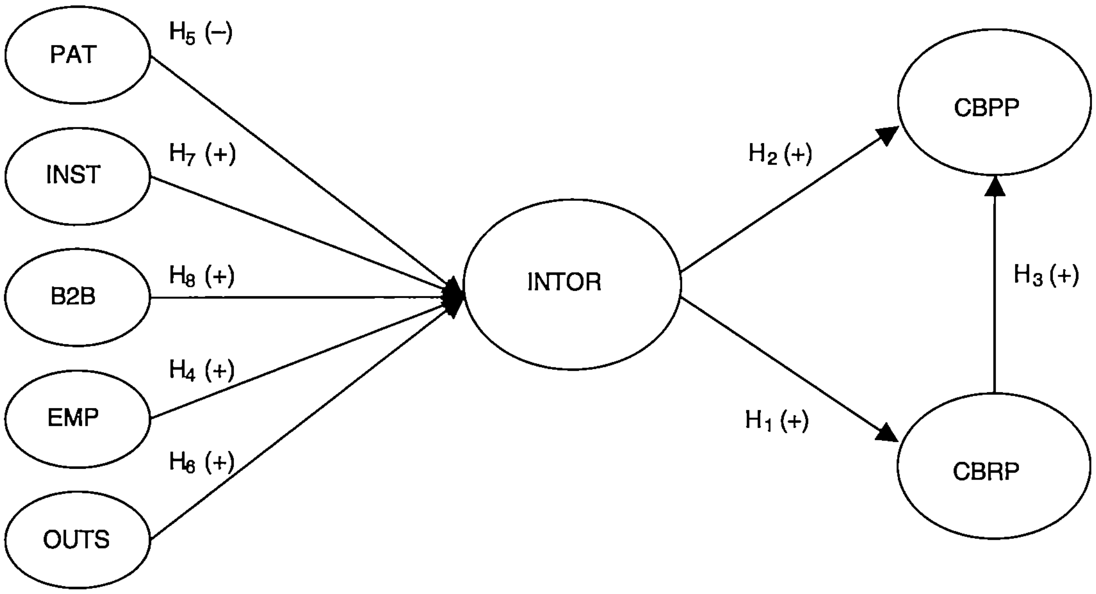
In our hypothesized model, the focal or central variable is INTOR because it performs a mediating role between the antecedents and the consequence constructs. In other words, the hypothesized model does not have direct paths from the antecedents (EMP, PAT, B2B, INST, and OUTS) to the consequence constructs (CBRP and CBPP). In the rival model, we allowed the five antecedent variables and INTOR to have direct effects on CBRP and CBPP. Thus, in the rival model, INTOR is not a mediating construct. We compared the hypothesized structural model with the rival model. This helps us test the nomological status of the focal variable (De Wulf, Odekerken-Schroder, and Iacobucci 2001; Morgan and Hunt 1994). The rival model (number of distinct parameters to be estimated = 67) was less parsimonious than the hypothesized model (number of distinct parameters to be estimated = 57): $\chi^2 = 287.17$, d.f. = 143; CFI = .94; GFI = .81; TLI = .92; IFI = .94; and SRMR = .07. Because the two models use the exact same covariance structure as the input, and thus are nested, we compared the two models using the chi-square difference test. The test results ($\Delta\chi^2 = 14.95$, Δd.f. = 10, $p > .05$) indicated that the rival model did not explain the covariance structure any better than the hypothesized model. We also compared the two models on the basis of the percentage of the model's estimated paths that are statistically significant (Morgan and Hunt 1994). The percentage of estimated paths supported in the hypothesized model (18/20 = 90%) was greater than the percentage of estimated paths supported in the rival model (14/25 = 56%). Furthermore, when we examined only the paths between the latent constructs and ignored the paths

TABLE 3
Common Method Bias Analysis

	OUTS	EMP	INST	CBPP	INTOR	PAT	COMPIN (MV)
EMP	.61 .53 .51						
INST	.80 .76 .75	.51 .40 .38					
CBPP	.80 .76 .75	.68 .61 .60	.65 .57 .56				
INTOR	.88 .86 .85	.74 .69 .68	.78 .73 .72	.86 .83 .82			
PAT	.58 .49 .48	.62 .54 .53	.40 .28 .25	.75 .70 .69	.73 .67 .65		
COMPIN (MV)	*.05* *−.16* −.20	*.05* *−.15* −.20	*−.14* −.38 −.43	*.03* *−.18* −.22	*.12* *−.07* *−.11*	*.11* *−.09* *−.13*	
CBRP	.76 .71 .70	.66 .59 .57	.65 .57 .56	.78 .73 .72	.84 .80 .79	.61 .53 .51	*.17 (MV)*

Notes: MV = marker variable. All correlations are significant at $p < .05$, except for values in italics. The first value in each cell is the correlation between the constructs, the second value is the correlation corrected for method bias, and the third value is the correlation value for a 95% sensitivity analysis. For this analysis, we reverse-scored the variable PAT.

FIGURE 3
The Hypothesized Structural Model



representing the loading of items onto their respective factors, the ratio of supported paths to hypothesized paths was 6 of 8 (75%) in our hypothesized model, and the ratio of significant paths to examined paths was only 2 of 13 (15.38%) in the rival model.

Thus, we prefer the more parsimonious hypothesized model to the rival model. This result also implies that INTOR holds a central nomological status and therefore is a key construct in explaining both a firm's CBRP and CBPP.

Moderator Regression Analyses

To test H_9, we carried out a moderator regression analysis to determine whether a significant interaction effect existed between INTOR and CIC, with CBPP as the dependent variable. To test H_{10}, we carried out a separate moderator regression analysis to determine whether a significant interaction effect existed between INTOR and COMPIN, with CBRP as the dependent variable. We obtained scores for the

TABLE 4
Results of the Hypothesis Testing

A: Structural Model: H_1–H_8

Construct	Direction	Construct	Estimate	Standardized Estimate	SE	t-Value	p	Hypothesis	Conclusion
CBRP	←	INTOR	1.19	.84	.10	11.80	.00	1	Supported
CBPP	←	INTOR	.77	.77	.12	6.38	.00	2	Supported
CBPP	←	CBRP	.09	.13	.08	1.08	.28	3	Not supported
INTOR	←	PAT	–.22	–.27	.05	–4.48	.00	5	Supported
INTOR	←	EMP	.31	.20	.10	3.27	.00	4	Supported
INTOR	←	INST	.25	.21	.10	2.43	.02	7	Supported
INTOR	←	OUTS	.31	.42	.07	4.48	.00	6	Supported
INTOR	←	B2B	.07	.03	.10	.75	.46	8	Not supported

B: Moderator Regression: H_9

	CBPP					
Dependent Variable	**β**	**t-Value**	**VIF**	**F**	**p**	**R^2**
Main effects–only model				205.54	.00	.90
INTOR	.55	15.56	1.11		.00	
CIC	.32	7.34	1.11		.00	
Moderator effects model				142.83	.00	.91
INTOR	.55	15.80	1.11		.00	
CIC	.33	7.58	1.12		.00	
INTOR × CIC	.06	1.97	1.01		.05	

Notes: VIF = variance inflation factor.

CBPP and CBRP variables for this analysis by averaging the respondent ratings across the respective indicators. The CIC moderator was a single-item measure, and we computed COMPIN as the average of two items (see the Appendix). We followed the procedure outlined by Aiken and West (1991) and used by other researchers (e.g., Jones and Reynolds 2006) to carry out the two moderator regression analyses. We mean-centered the independent variables INTOR, CIC, and COMPIN to reduce multicollinearity. The variance inflation factors were below 2 in the regression involving CIC and below 3.5 in the regression involving COMPIN. This indicated that multicollinearity was not an issue.

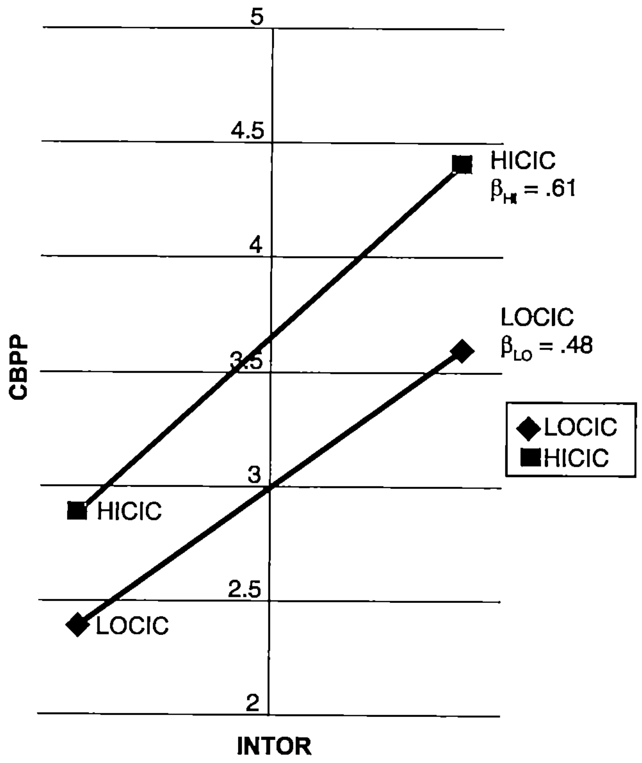
We ran a regression involving only the main terms INTOR and CIC on the dependent variable CBPP. We then added the product term INTOR × CIC as a predictor variable and repeated the regression. We report the results of the regression analyses in Table 4, Panel B.

In the main effects–only model, the main effect of INTOR on CBPP ($\beta = .55, p < .01$) is consistent with our prior analysis. The main effect of CIC on CBPP is also significant ($\beta = .32, p < .01$). Although we did not hypothesize this main effect, it seems reasonable to expect that firms with a high percentage of proactive customers will exhibit superior CBPP. This is also consistent with the findings of Bowman and Narayandas (2001), who report that customers who initiate contact with manufacturers are among a brand's most loyal customers in terms of share-of-category requirements. The moderator effects model had a marginally higher R-square. However, interaction effects rarely increase explained variance, and the focus in models with interaction effects is on understanding the observed relationships, not on predicting the dependent variable (Aiken and West 1991; Jones and Reynolds 2006). Again, the two main effects are significant. The product term (INTOR × CIC) is also significant ($\beta = .06, p = .05$). This result suggests that the variable CIC is a quasi moderator (Sharma, Durand, and Gur-Arie 1981; Voss and Voss 2000) of the relationship between INTOR and CBPP. Figure 4 shows a plot of predicted values created using low and high values of INTOR and CIC.

The low values reflect one standard deviation below the mean, and the high values reflect one standard deviation above the mean for each of the two variables INTOR and CIC (Aiken and West 1991; Jones and Reynolds 2006). The plot indicates that the slope for the high CIC value is higher than the slope for the low CIC value. The slopes ($\beta_{HI} = .61$, and $\beta_{LO} = .48$) are significantly different because the regression coefficient for the interaction term (CIC × INTOR) is significant (Aiken and West 1991). The results of the regression and the plot suggest that the effect of INTOR on CBPP performance is enhanced in the presence of higher CICs. Thus, H_9 is supported.

We followed the same procedure to examine whether COMPIN moderates the effect of INTOR on CBRP. The product term (INTOR × COMPIN) is not significant ($\beta = .05, p = .50$). Thus, H_{10} is not supported. Thus, we conclude that COMPIN does not moderate the effect of INTOR on CBRP.

FIGURE 4
A Plot of Predicted Values of CBPP for Low CICs (LOCIC) and High CICs (HICIC) from the Moderator Regression Analysis



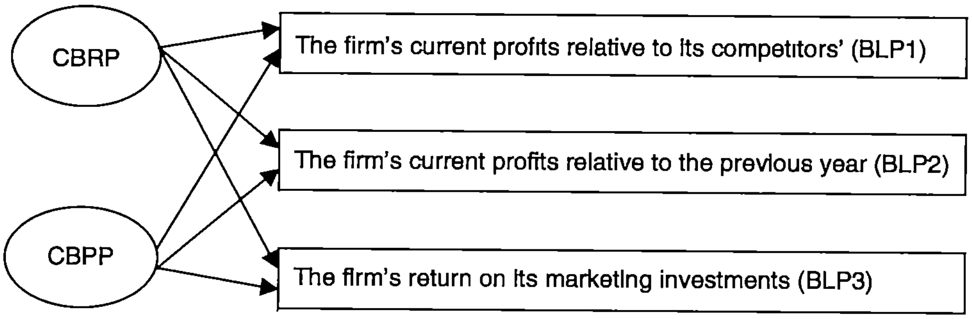
Extended Analysis: Customer-Based Performance and Firm-Level Performance

In the analysis we reported previously, the relationship between CBRP and CBPP was not supported. The literature on relationship marketing suggests that superior relational performance should lead to superior aggregate firm-level business performance (Boulding et al. 2005). However, researchers have also found that satisfaction and loyalty do not always directly translate into profits. Customer satisfaction is a helpful but not a sufficient condition for maintaining customer loyalty, and in many industries, it has become virtually a competitive requirement (Peppers and Rogers 2004). A focus on customer satisfaction may not be effective for all kinds of customers (Garbarino and Johnson 1999). In addition, not all loyal customers are necessarily profitable (Reinartz and Kumar 2003).

We conducted a further analysis of the impact of CBRP and CBPP on aggregate firm-level business performance using three single-item self-report measures: (1) the firm's current profits relative to competitors' (BLP1), (2) the firm's current profits relative to the previous year (BLP2), and (3) the firm's return on its marketing investments (BLP3). We retained all the variables and paths of the hypothesized structural model and extended the model by adding three paths each from CBRP and CBPP to the three single-item measures BLP1, BLP2, and BLP3.[5] Thus, we specified a total of six additional paths (see Figure 5).

Modification indexes obtained from an initial test of the model suggested that the model fit would improve if we allowed the errors between BLP1 and BLP2 to covary. This modification was theoretically grounded because BLP1 (i.e., profits relative to competitors') and BLP2 (i.e., profits relative to last year) are likely to have common causal variables not included in the model. Therefore, we respeci-

FIGURE 5
Customer-Based Performance and Aggregate Business-Level Performance


CBRP
CBPP
The firm's current profits relative to its competitors' (BLP1)
The firm's current profits relative to the previous year (BLP2)
The firm's return on its marketing investments (BLP3)

Notes: Although we depict only six additional paths here, the actual model includes the antecedents of INTOR, the INTOR construct, and the paths representing H_1–H_8

---

[5] The extended analysis is a logical inquiry resulting from the findings in the first stage of the analysis. Because the findings in the first stage indicated that the path between CBRP and CBPP was not significant, we could explain this somewhat counterintuitive result by adding the firm-level performance variables in the second stage of the analysis. The two-stage approach also ensured that the hypotheses testing was carried out on a parsimonious one-stage mediation model rather than on a more complex two-stage mediation model. The two-stage analysis also helps us establish the nomological validity of the interaction orientation construct by comparing the first-stage model (consisting of one mediating effect) with a rival model (consisting of no mediating effect). We then added a second stage of mediation to the model by including three paths each from CBPP and CBRP to BLP1, BLP2, and BLP3.

fied the model to include the covariance between the error terms of BLP1 and BLP2.

Results of the Extended Analysis

The respecified model fit statistics were as follows: χ^2 = 484.71, d.f. = 209; CFI = .90; GFI = .75; TLI = .88; IFI = .90; and SRMR = .07. The estimated path coefficients of the six additional paths in the extended structural model appear in Table 5. The estimated coefficients for the paths common with the originally hypothesized model were similar in sign and significance, and thus we do not report them for this extended model.

As we expected, the paths from CBPP to BLP1 (β = .717, $p < .01$), BLP2 (β = .53, $p < .01$), and BLP3 (β = .517, $p < .01$) were significant. Two of the three paths from CBRP to BLP2 (β = .20, $p < .05$) and BLP3 (β = .21, $p < .05$) were also significant, thus reaffirming the belief that superior CBRP would lead to superior aggregate BLP. The two paths that support the CBRP–BLP link pertain to aggregate BLP measures that imply time-lag effects. For example, BLP2 captured information on improvements in profits over the previous year, and BLP3 captured information on return on marketing investments. The dependent variable in the path that was not significant, BLP1 (β = .05, $p > .05$), captured information on current profits relative to the competition. This result suggests that superior relational performance translates into superior aggregate BLP over time. Thus, attempts to enhance the INTOR of a firm lead to improved CBRP. In turn, improved CBRP leads to superior aggregate BLP in the long run. Our study was cross-sectional in design, and temporal effects cannot be inferred conclusively. Nevertheless, this additional analysis corroborates the existence of the CBRP–BLP linkage.

Although it is believed that the impact of strategic initiatives should be examined against objective measures of performance, prior research has helped establish strong links between perceptual measures of performance and objective measures (e.g., Jayachandran and Varadarajan 2006; Naman and Slevin 1993). Therefore, we expect our findings to hold for objective measures of performance. However, we proceeded to obtain from published sources an objective measure of aggregate BLP (OBLP2) that captured the growth in net income over the previous year for 94 firms in our database. This measure corresponds to the perceptual measure BLP2 we obtained through the survey. A correlation of .85 between BLP2 and OBLP2 indicates a close association between these two measures in our data. We repeated the extended analysis by substituting OBLP2 for BLP2, and the results we obtained were similar to the results from the extended analysis model that contained the BLP2 perceptual measure.

Discussion

We examine the contributions of our study in the context of the recent article by Boulding and colleagues (2005) that captures the state of CRM research and lays out a road map for the future. Customer relationship management represents the evolution and integration of marketing ideas with technology (Boulding et al. 2005). Our contribution is in identifying, distilling, and specifying an emergent marketing belief (customer concept) and the technological processes (interaction response capacity) and practices (customer empowerment and customer value management) that a firm and its managers must adopt to carry this evolution forward. Payne and Frow (2005) point out that CRM has an array of metrics that are outcome measures but few measures that drive value creation. The four elements of interaction orientation that we offer are precise and actionable measures that can indicate where a firm is headed in its efforts to manage customer relationships. Through our study, we establish that an interaction orientation leads to superior performance outcomes. This provides further empirical support for Boulding and colleagues' proposition that CRM practices enhance firm performance. Boulding and colleagues identify that dual creation of value is at the heart of CRM. We offer interaction response capacity and customer empowerment, with its focus on customer-to-customer interactivity, as the two elements of interaction orientation that reflect dual creation of value. Boulding and colleagues also emphasize the need for CRM research that leads to generalizable results. The use of multiple respondents across a large cross-section of firms in varied industries helps us achieve this to a great extent.

Capabilities are complex bundles of skills and accumulated knowledge exercised through organizational processes that enable firms to coordinate activities and make use of their assets (Day 1994). Organizational resources determine the efficacy of the strategic response (Lee and Grewal 2004). An organizational capability is a knowledge-based resource (Capron and Hulland 1999) and refers to the processes and routines that a firm performs well (Slotegraaf, Moorman, and Inman 2003). Because the interaction orientation construct measures the extent to which a firm has

TABLE 5
Estimates of Six Additional Path Coefficients (Extended Structural Model with Aggregate Business-Level Performance Measures)

Dependent Variable	Direction	Construct	Estimate	Standardized Estimate	SE	t-Value	*p*	Conclusion
BLP1	←	CBPP	.71	.73	.13	5.21	.00	Supported
BLP2	←	CBPP	.53	.54	.13	4.09	.00	Supported
BLP3	←	CBPP	.51	.53	.13	3.95	.00	Supported
BLP1	←	CBRP	.05	.08	.09	.61	.54	Not supported
BLP2	←	CBRP	.20	.30	.08	2.32	.02	Supported
BLP3	←	CBRP	.21	.31	.09	2.35	.02	Supported

developed relevant processes and routine practices based on dynamically accumulating and using individual customer knowledge, it is a measure of a knowledge-based resource. We developed interaction orientation as a composite construct encompassing an underlying belief and the processes and practices that supplement this belief. When these processes and practices that constitute an interaction orientation are implemented by a firm, they develop into an inimitable complex bundle of skills that lead to continued superior firm performance relative to competition. Thus, an interaction orientation, as we conceptualize it, can be viewed as a composite of the underlying belief, the customer concept, and the concurrent capabilities that a firm develops.

Managerial Implications

Over the years, firms that succeeded by adopting a product orientation gave way to firms with a sales orientation, which in turn gave way to firms with a market orientation. Today, firms are learning to succeed by adopting an interaction orientation. Firms still need to produce superior products, sell smarter, and understand the market as a whole, but the ability of firms to orient themselves to interact successfully with their individual customers will differentiate firms in the future. Interaction orientation is different from market orientation in that (1) firms that are interaction oriented believe that the individual customer and not a market is the unit of analysis, (2) marketing activities are conducted with the customer rather than for the customer, (3) customer-to-customer linkages are strategically important to the firm, and (4) monetary accountability for marketing actions is entrenched in practice.

The findings from our study have several implications for firms. Firms dependent on protected patents may not reap continued benefits. Firms must adopt customer-based performance metrics and institute employee rewards and incentives based on these metrics, instead of using aggregate-level measures, such as sales and market share growth, to evaluate marketing performance. Firms must develop their outsourcing capabilities by adopting sound contracting procedures so that they can offer diverse products and services their individual customers desire. An interaction orientation is relevant not only to business-to-business industries but also to business-to-consumer industries. Firms should focus on building an interaction orientation, regardless of whether the competitive intensity is high or low. The proposed INTOR scale could also be used to observe the link between the level of interaction orientation and performance variables as firms introduce various customer-directed programs.

The results from this study imply that the managerial benefits of adopting an interaction orientation are that (1) the firm is able to attract and retain the most valuable customers, (2) the firm's customers develop into a skilled resource for the firm, (3) the firm inoculates its customers from competitors by instilling a heightened sense of ownership of the firm among its customers, (4) the firm develops a dynamically shifting portfolio of products and services, (5) the firm develops the ability to foresee customer responses, and (6) the firm exhibits superior aggregate business-level performance because interaction orientation results in the firm dynamically maximizing the profit function at every stage of activity across all customers. As a corollary, not adopting a sufficient degree of interaction orientation could lead to (1) a large base of unprofitable customers with an obligation for the firm to continue serving them, (2) uncontrolled proliferation of negative word of mouth, (3) customers directing competitors' attention to the firm's vulnerable areas, (4) an inability to offer a relevant range of products and services, (5) an inability to plan into the future, and (6) poor marketing accountability, leading to lower profits and return on marketing investments.

Wells Fargo is an example of an organization that believes in the customer concept. It has continuously made investments in technologies that help maintain real-time visibility of the firm's customers. According to Danny Peltz, executive vice president of Wells Fargo, these efforts have resulted in increased agility versus competition and an upward trend in revenues, transaction volume, and services per customer (BEA Systems Inc. 2006). An example of a firm that illustrates the successful adoption of an interaction response capacity is Boeing. Boeing's Rapid Response Center handles technical problems for its airline customers that arise out of normal business hours by using interactive video, Internet, and telephone systems that seamlessly allow access to people and data across multiple sites (www.myboeingfleet.com). IBM supports Linux to provide more choices to its customers. Customers appreciate this freedom from expensive licensed software. This action demonstrates the recognition of customer empowerment as a strategic activity by IBM. IBM has also developed sophisticated techniques to assess future customer value of its institutional customers. This focus on customer value management has helped IBM improve profitability in the U.S. market over the past few years (Kumar et al. 2007).

New entrants into the business-to-consumer industries have also established themselves by adopting an interaction orientation. Whereas record labels such as EMI and Sony were successful in the past, Apple, with its iPod and iTunes combination, has rejuvenated the music industry through an emphasis on individual customer-level interactions and customer-to-customer interactions.

Our study shows that there is considerable variation in the level of interaction orientation between firms, and firms that have nurtured an interaction orientation exhibit superior outcome in terms of a range of firm performance measures. Thus, firms that focus on the four components of an interaction orientation can develop a capability that leads to marketing success. Therefore, we conclude that an interaction orientation, in the current paradigm of interactive marketing, is a rare source of competitive advantage. However, with time, a new paradigm of marketing may well evolve.

Limitations and Suggestions for Further Research

Although our study establishes that organizations embracing an interaction orientation perform better, longitudinal studies would help understand whether the nature of the benefits accrued is indeed long-term. Linkages within the dimensions of interaction orientation, similar to the effect of the strengthening of the belief in the customer concept on the development of superior interaction response capacities, customer value management, and customer empowerment practices over time, could also be examined with longitudinal data. Richer insights might be available if future studies are conducted across many firms within an industry. Factors such as entrepreneurship, innovativeness, capacity to innovate, learning orientation, and market orientation are all known to affect a firm's positional advantage (Hult and Ketchen 2001; Hurley and Hult 1998). Future empirical studies could examine the relative contribution of an interaction orientation and the collective contribution of all these factors together to a firm's positional advantage.

Our study examines two of many moderating variables of the impact of an interaction orientation on firm performance. Market turbulence, technological turbulence, and absorptive capacity could be examined in future studies. Although we examined the moderating effect of customer-initiated contacts, we did not differentiate between contacts that arise as a result of complaining behaviors and other positive kinds of contacts.

Appendix

Note 1: The word "customer" refers to direct institutional customers for business-to-business firms, to end consumers for business-to-consumer firms, and to retailers for business-to-retail firms. Note 2: Unless mentioned otherwise, questions are to be answered on a five-point scale, where 1 = "strongly agree" and 5 = "strongly disagree." Please circle the appropriate number on the scale provided against the question.

Questionnaire Items to Measure Interaction Orientation

Belief in the Customer Concept (CC)

1. This firm believes that each customer cannot be satisfied with the same set of products and services. (CC1)
2. This firm consciously seeks to identify and acquire new customers individually. (CC2)
3. This firm believes that customers' reactions to marketing action should be observed at the individual customer level. (CC3)

Interaction Response Capacity (IRC)

4. This firm has systems in place that record each customer's transactions. (IRC1)
5. This firm can identify all transactions pertaining to each individual customer. (IRC2)
6. This firm analyzes previous consumer transactions at the individual customer level to predict future transactions from that customer. (IRC3)
7. In this firm, all customer interfaces possess transaction information on individual customers at all times. (IRC4)

Customer Empowerment (CE)

8. This firm encourages customers to share opinions of its products or services with the firm. (CE1)
9. This firm encourages customers to share opinions of its products or services with other customers. (CE2)
10. This firm encourages customers to participate interactively in designing products and services. (CE3)

Customer Value Management (CVM)

11 This firm has an excellent idea of what each individual customer has been contributing to its profits. (CVM1)
12. This firm predicts what each individual customer will contribute to its profits in the future. (CVM2)
13. This firm computes the revenue generated as a result of every marketing action directed at an individual customer. (CVM3)

Questionnaire Items to Measure the Consequences of Interaction Orientation

Customer-Based Profit Performance (CBPP)

Identification of profitable customers:

1. Customers who this firm identifies as potentially profitable turn out to be profitable in the long run. (CBPP1)

Efficiency of customer acquisition and retention:

2. A larger proportion of acquired customers remain profitable in the long run for this firm as compared to its competitors. (CBPP2)

Conversion of unprofitable customers to profitable ones:

3. The number of customers who were unprofitable last year and became profitable this year for this firm is greater than the number of customers who were profitable last year but became unprofitable this year. (CBPP3)

Customer-Based Relational Performance (CBRP)

Customer satisfaction:

4. The overall satisfaction level of our customers is higher than the satisfaction levels of these customers with our competing firms. (CBRP1)

Customer ownership:

5. This firm's customers are interested in the financial well-being of the firm. (CBRP2)

Customer word of mouth:

6. A higher percentage of our new customers come to us because of referrals from our existing customers, relative to our competitors. (CBRP3)

Aggregate Business-Level Performance (1 = "much lower," and 5 = "much higher")

Relative profits (adapted from Deshpandé, Farley, and Webster 1993)

7. Relative to this firm's main competitors, currently our profits are ... (BLP1)
8. Relative to last year, this firm's profits are ... (BLP2)

Marketing return on investment (adapted from Deshpandé and Farley 1998; Narver, Jacobson, and Slater 1999):

9. Relative to our firm's main competitors, our marketing investments result in returns that are ... (BLP3)

Questionnaire Items to Measure the Antecedents of Interaction Orientation

Dependence on Trademarks (PAT)

1. This firm closely guards its patents and trademarks. (PAT1)
2. This firm is able to stay ahead of its competition because of the legal protection it enjoys on its trademarks or patents. (PAT2)

Normative Institutional Pressure (INST) (adapted from Wu, Mahajan, and Balasubramanian 2003)

3. A large number of competitors of this firm have adopted interactive technologies. (INST1)
4. In this industry, firms that do not readily adopt new interactive technologies will be left behind. (INST2)

Employee Reward System (EMP) (adapted from Jaworski and Kohli (1993)

5. In this firm, employee compensation depends on individual performance on customer measures like customer acquisition, customer retention, and customer win-backs. (EMP1)
6. In this firm, employee compensation is linked to firm-level sales. (EMP2)
7. In this firm, employee compensation is linked to achievement of employee-level sales quota. (EMP3)

Outsourcing Expertise (OUTS)

8. This firm is excellent at managing its outsourcing activities. (OUTS1)
9. The quality of this firm's outsourcing partners is ... (OUTS2) (1 = "extremely low," and 5 = "extremely high")

Questionnaire Items to Measure Moderator Variables

Competitive Intensity (COMPIN) (adapted from Jaworski and Kohli 1993)

1. In our business, any product or service that one competitor can offer, others can match readily.
2. Our competitors are relatively weak.

Customer Initiated Contacts (CIC)

3. The number of customers who initiated communications with the firm this year (expressed as a percentage of the total number of customers that the firm caters to) is ________%. (Enter percentage rounded off to the nearest whole number in the box provided below; e.g., 10)

REFERENCES

Aiken, Leona S. and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage Publications.

Anderson, Eugene W., Claes Fornell, and Sanal K. Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value," *Journal of Marketing*, 68 (October), 172–85.

Anderson, James C. and David W. Gerbing (1988), "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (3), 411–23.

Atuahene-Gima, Kwaku and Janet Y. Murray (2004), "Antecedents and Outcomes of Marketing Strategy Comprehensiveness," *Journal of Marketing*, 68 (October), 33–46.

BEA Systems Inc. (2006), "Wells Fargo: Enhancing Service Delivery and Growing Revenue Through Shared Services," customer case study, (accessed October 12, 2006), [available at http://www.bea.com/content/news_events/white_papers/BEA_WellsFargo_ss.pdf].

Bharadwaj, Sundar G., P. Rajan Varadarajan, and John Fahy (1993), "Sustainable Competitive Advantage in Service Industries: A Conceptual Model and Research Propositions," *Journal of Marketing*, 57 (October), 83–99.

Boulding, William, Richard Staelin, Michael Ehret, and Wesley J. Johnston (2005), "A Customer Relationship Management Roadmap: What Is Known, Potential Pitfalls, and Where to Go," *Journal of Marketing*, 69 (October), 155–66.

Bowman, Douglas and Das Narayandas (2001), "Managing Customer-Initiated Contacts with Manufacturers: The Impact of Share of Category Requirements and Word-of-Mouth Behavior," *Journal of Marketing Research*, 38 (August), 281–97.

Brown, Tom J., Thomas E. Barry, Peter A. Dacin, and Richard F. Gunst (2005), "Spreading the Word: Investigating Antecedents of Consumers' Positive Word-of-Mouth Intentions and Behaviors in a Retailing Context," *Journal of the Academy of Marketing Science*, 33 (2), 123–38.

Burke, Michael J. and William P. Dunlap (2002), "Estimating Interrater Agreement with the Average Deviation Index: A User's Guide," *Organizational Research Methods*, 5 (2), 159–72.

Capron, Laurence and John Hulland (1999), "Redeployment of Brands, Sales Forces, and General Marketing Management Expertise Following Horizontal Acquisitions: A Resource-Based View," *Journal of Marketing*, 63 (April), 41–54.

Churchill, Gilbert A., Jr. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16 (February), 64–74.

Coviello, Nicole E., Roderick J. Brodie, Peter J. Danaher, and Wesley J. Johnston (2002), "How Firms Relate to Their Markets: An Empirical Examination of Contemporary Marketing Practices," *Journal of Marketing*, 66 (July), 33–46.

Day, George S. (1994), "The Capabilities of Market-Driven Organizations," *Journal of Marketing*, 58 (October), 37–52.

——— (1999), *The Market Driven Organization: Understanding, Attracting, and Keeping Valuable Customers*. New York: The Free Press.

——— (2000), "Managing Market Relationships," *Journal of the Academy of Marketing Science*, 28 (1), 24–30.

——— and David B. Montgomery (1999), "Charting New Directions for Marketing," *Journal of Marketing*, 63 (October), 3–13.

——— and Christophe Van den Bulte (2002), "Superiority in Customer Relationship Management: Consequences for Competitive Advantage and Performance," Marketing Science Institute Working Paper No. 02-123.

De Wulf, Kristof, Gaby Odekerken-Schroder, and Dawn Iacobucci (2001), "Investments in Consumer Relationships: A Cross-Country and Cross-Industry Exploration," *Journal of Marketing*, 65 (October), 33–50.

Deshpandé, Rohit and John U. Farley (1998), "Measuring Market Orientation: Generalization and Synthesis," *Journal of Market-Focused Management*, 2 (3), 213–32.

———, ———, and Frederick E. Webster Jr. (1993), "Corporate Culture, Customer Orientation, and Innovativeness," *Journal of Marketing*, 57 (January), 23–37.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–51.

Garbarino, Ellen and Mark S. Johnson (1999), "The Different Roles of Satisfaction, Trust, and Commitment in Customer Relationships," *Journal of Marketing*, 63 (April), 70–87.

Gerbing, David W. and James C. Anderson (1988), "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment," *Journal of Marketing Research*, 25 (May), 186–93.

Grewal, Rajdeep, James M. Comer, and Raj Mehta (2001), "An Investigation into the Antecedents of Organizational Participation in Business-to-Business Electronic Markets," *Journal of Marketing*, 65 (July), 17–33.

———, Joseph A. Cote, and Hans Baumgartner (2004), "Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing," *Marketing Science*, 23 (4), 519–29.

Grönroos, Christian (1994), "From Marketing Mix to Relationship Marketing: Towards a Paradigm Shift in Marketing," *Management Decision*, 32 (2), 4–21.

Hoekstra, Janny C., Peter S.H. Leeflang, and Dick R. Wittink (1999), "The Customer Concept: The Basis for a New Marketing Paradigm," *Journal of Market-Focused Management*, 4 (1), 43–76.

Hogan, John E., Katherine N. Lemon, and Barak Libai (2003), "What Is the True Value of a Lost Customer?" *Journal of Service Research*, 5 (3), 196–208.

Homburg, Christian and Christian Pflesser (2000), "A Multiple-Layer Model of Market-Oriented Organizational Culture: Measurement Issues and Performance Outcomes," *Journal of Marketing Research*, 37 (November), 449–62.

Hu, Li-Tze and Peter M. Bentler (1999), "Cut-Off Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives," *Structural Equation Modeling*, 6 (1), 1–55.

Hult, G. Tomas M. and David J. Ketchen (2001), "Does Market Orientation Matter? A Test of the Relationship Between Positional Advantage and Performance," *Strategic Management Journal*, 22 (9), 899–906.

Hunt, Shelby D. and Robert M. Morgan (1995), "The Comparative Advantage Theory of Competition," *Journal of Marketing*, 59 (April), 1–15.

Hurley, Robert F. and G. Tomas M. Hult (1998), "Innovation, Market Orientation, and Organizational Learning: An Integration and Empirical Examination," *Journal of Marketing*, 62 (July), 42–54.

Jaworski, Bernard J. and Ajay K. Kohli (1993), "Market Orientation: Antecedents and Consequences," *Journal of Marketing*, 57 (July), 53–70.

Jayachandran, Satish, Subhash Sharma, Peter Kaufman, and Pushkala Raman (2005), "The Role of Relational Information Processes and Technology Use in Customer Relationship Management," *Journal of Marketing*, 69 (October), 177–92.

——— and Rajan Varadarajan (2006), "Does Success Diminish Competitive Responsiveness? Reconciling Conflicting Perspectives," *Journal of the Academy of Marketing Science*, 34 (3), 284–95.

Jones, Michael A. and Kristy E. Reynolds (2006), "The Role of Retailer Interest on Shopping Behavior," *Journal of Retailing*, 82 (2), 115–26.

Kandemir, Destan, Attila Yaprak, and S. Tamer Cavusgil (2006), "Alliance Orientation: Conceptualization, Measurement, and Impact on Market Performance," *Journal of the Academy of Marketing Science*, 34 (3), 324–41.

King, William R. (2004), "Outsourcing and the Future of It," *Information Systems Management*, 21 (4), 83–84.

Kirca, Ahmet H., Satish Jayachandran, and William O. Bearden (2005), "Market Orientation: A Meta-Analytic Review and Assessment of Its Antecedents and Impact on Performance," *Journal of Marketing*, 69 (April), 24–41.

Kline, Rex B. (2005), *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press.

Kohli, Ajay K., Bernard J. Jaworski, and Ajith Kumar (1993), "MARKOR: A Measure of Market Orientation," *Journal of Marketing Research*, 30 (November), 467–77.

Kumar, Nirmalya, Louis W. Stern, and James Anderson (1993), "Conducting Interorganizational Research Using Key Informants," *Academy of Management Journal*, 36 (6), 1633–51.

Kumar, V., J. Andrew Petersen, and Robert P. Leone (2006), "The Power of Customer Advocacy," in *AMA Winter Educators' Conference Proceedings*, Vol. 17 (CD-ROM), Jean L. Johnson and John Hulland, eds. Chicago: American Marketing Association, 81–82.

———, Girish Ramani, and Timothy Bohling (2004), "Customer Lifetime Value Approaches and Best Practice Applications," *Journal of Interactive Marketing*, 18 (3), 60–72.

——— and Werner Reinartz (2006), *Customer Relationship Management. A Databased Approach*. Hoboken, NJ: John Wiley & Sons.

———, R. Venkatesan, T. Bohling, and D. Beckmann, (2007), "The Power of CLV: Managing Customer Value at IBM," *Marketing Science*, forthcoming.

Lee, Ruby P. and Rajdeep Grewal (2004), "Strategic Responses to New Technologies and Their Impact on Firm Performance," *Journal of Marketing*, 68 (October), 157–71.

Lindell, Michael K. and David J. Whitney (2001), "Accounting for Common Method Variance in Cross-Sectional Research Designs," *Journal of Applied Psychology*, 86 (1), 114–21.

Lohr, Steve (2005), "I.B.M. to Give Free Access to 500 Patents," *The New York Times*, (January 11), (accessed September 13, 2007), [available at http://www.nytimes.com/2005/01/11/technology/11soft.html?ex=1189828800&en=d7d8733c980f561b&ei=5070].

Matsuno, Ken, John T. Mentzer, and Aysegul Ozsomer (2002), "The Effects of Entrepreneurial Proclivity and Market Orientation on Business Performance," *Journal of Marketing*, 66 (July), 18–33.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment–Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–38

Mulhern, Francis J. (1999), "Customer Profitability Analysis: Measurement, Concentration, and Research Directions," *Journal of Interactive Marketing*, 13 (1), 25–40.

Naman, John L. and Dennis P. Slevin (1993), "Entrepreneurship and the Concept of Fit: A Model and Empirical Tests," *Strategic Management Journal*, 14 (2), 137–53.

Narver, John C., Robert L Jacobson, and Stanley F. Slater (1999), "Market Orientation and Business Performance: An Analysis of Panel Data," in *Developing a Market Orientation*, Rohit Deshpandé, ed. Thousand Oaks, CA: Sage Publications, 195–216.

Newell, Frederick (2003), *Why CRM Doesn't Work*? Princeton, NJ: Bloomberg Press.

Noble, Charles H., Rajiv K. Sinha, and Ajith Kumar (2002), "Market Orientation and Alternative Strategic Orientations: A Longitudinal Assessment of Performance Implications," *Journal of Marketing*, 66 (October), 25–39.

Nunnally, Jum C. (1978), *Psychometric Theory*. New York: McGraw-Hill.

Payne, Adrian and Pennie Frow (2005), "A Strategic Framework for Customer Relationship Management," *Journal of Marketing*, 69 (October), 167–76.

Peppers, Don and Martha Rogers (2004), *Managing Customer Relationships: A Strategic Framework*. Hoboken, NJ· John Wiley & Sons.

Prahalad, C.K. and Venkatram Ramaswamy (2004), *The Future of Competition: Co-creating Unique Value with Customers.* Boston: Harvard Business School Press.

Rayport, Jeffrey F. and Bernard J. Jaworski (2005), *Best Face Forward.* Boston: Harvard Business School Press.

Reichheld, Fredrick (2006), *The Ultimate Question.* Boston: Harvard Business School Press.

Reinartz, Werner J., Manfred Krafft, and Wayne D. Hoyer (2004), "The Customer Relationship Management Process: Its Measurement and Impact on Performance," *Journal of Marketing Research*, 41 (August), 293–305.

——— and V. Kumar (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (January), 77–99.

Reitzig, Markus (2004), "Strategic Management of Intellectual Property," *MIT Sloan Management Review*, 45 (3), 35–40.

Sharma, Subhash, Richard M. Durand, and Oded Gur-Arie (1981), "Identification and Analysis of Moderator Variables," *Journal of Marketing Research*, 18 (August), 291–300.

Sheth, Jagdish N., Rajendra S. Sisodia, and Arun Sharma (2000), "The Antecedents and Consequences of Customer-Centric Marketing," *Journal of the Academy of Marketing Science*, 28 (1), 55–66.

Slotegraaf, Rebecca J., Christine Moorman, and J. Jeffrey Inman (2003), "The Role of Firm Resources in Returns to Market Deployment," *Journal of Marketing Research*, 40 (August), 295–309.

Srinivasan, Srini S., Rolph Anderson, and Kishore Ponnavolu (2002), "Customer Loyalty in E-Commerce: An Exploration of Its Antecedents and Consequences," *Journal of Retailing*, 78 (1), 41–50.

Thomas, Jacquelyn S., Werner J. Reinartz, and V. Kumar (2004), "Getting the Most Out of All Your Customers," *Harvard Business Review*, 82 (7–8), 116–23.

Tsikriktsis, Nikos, Gianvito Lanzolla, and Mark Frohlich (2004), "Adoption of e-Processes by Service Firms: An Empirical Study of Antecedents," *Production and Operations Management*, 13 (3), 216–29.

Van Bruggen, Gerrit H., Gary L. Lilien, and Manish Kacker (2002), "Informants in Organizational Marketing Research: Why Use Multiple Informants and How to Aggregate Responses," *Journal of Marketing Research*, 39 (November), 469–78.

Varadarajan, P. Rajan and Manjit S. Yadav (2002), "Marketing Strategy and the Internet: An Organizing Framework," *Journal of the Academy of Marketing Science*, 30 (4), 296–312.

Venkatesan, Rajkumar and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (October), 106–125.

Voss, Glenn B. and Zannie Giraud Voss (2000), "Strategic Orientation and Firm Performance in an Artistic Environment," *Journal of Marketing*, 64 (January), 67–83.

Wu, Fang, Vijay Mahajan, and Sridhar Balasubramanian (2003), "An Analysis of E-Business Adoption and Its Impact on Business Performance," *Journal of the Academy of Marketing Science*, 31 (4), 425–47.

Yadav, Manjit S. and P. Rajan Varadarajan (2005), "Understanding Product Migration to the Electronic Marketplace: A Conceptual Framework," *Journal of Retailing*, 81 (2), 125–40.

Ruth N. Bolton, Katherine N. Lemon, & Peter C. Verhoef

Expanding Business-to-Business Customer Relationships: Modeling the Customer's Upgrade Decision

This article develops a model of a business customer's decision to upgrade service contracts conditional on the decision to renew the contract. It proposes that the firm's upgrade decision is influenced by (1) decision-maker perceptions of the relationship with the supplier, (2) contract-level experiences, and (3) interactions between firm- and contract-level variables. The authors model the firm's decision as a binary logit model with random parameters for the contract-level variables and fixed parameters for the firm-level variables. They estimate the model with data describing more than 2000 service contracts and find that decision-maker satisfaction, service quality, and price have a significant effect on the decision to upgrade; price and satisfaction also moderate the effect of service quality on the decision. Simulations indicate that modest improvements in service quality for a focal contract can have a relatively large, positive effect on the likelihood that the firm will upgrade. The results suggest that suppliers need to manage their firm relationships at both the individual contract level and the overall firm level. In addition, the results suggest specific windows of opportunity for suppliers when firms may be more likely to upgrade to higher-level service contracts.

Keywords: customer relationships, upgrades, service quality, satisfaction, retention

This article focuses on the customer's decision to upgrade a service contract. A product upgrade is a form of relationship expansion in which the customer purchases an expanded offering—a higher-price, augmented good or service (with higher service levels or additional features)—instead of repurchasing a low-price good or service (with lower service levels or fewer features) from the same supplier. Recently, researchers have developed predictive and normative models that are relevant to consumer service upgrades in industries such as airlines, banks, and theaters (e.g., Ngobo 2005). For example, Li, Sun, and Wilcox (2005) model a consumer's sequential acquisition of banking services on the basis of "maturity" (e.g., demographics, prior purchases). We have not been able to discover any theory-based models of the factors that influence a firm's decision to upgrade a service contract over time. This gap in the literature is rather remarkable because firms frequently buy fixed-price service contracts from suppliers in many industries, including computing, telecommunications, and information services; repair and maintenance services for engineering, medical, and/or other equipment; and services that support financial, health, or energy management systems. For example, General Electric, Hewlett-Packard, ServiceMaster, and Verizon are major suppliers of service contracts to firms. Service contracts are typically offered at different levels. For example, a "low-level" contract may promise reactive service (e.g., repairs), whereas a "high-level" service contract may promise proactive service (e.g., repairs and preventative maintenance). A large firm may purchase as many as 20–50 service contracts from a service supplier. An upgrade may occur on a given service contract whenever a firm has the opportunity to move (contractually) from a low level of service to a higher level of service.

Marketing academics and practitioners have long been interested in the nature of business-to-business (B2B) relationships (e.g., Dwyer, Schurr, and Oh 1987). An early study of B2B relationships found that the customer's overall assessment of the firm was a key determinant of his or her decision to continue to conduct business with the firm (Jackson 1985). A recent study by Cannon and Perreault (1999) shows how business customers' evaluations of supplier performance vary across different types of relationships. Research in services marketing has focused on cross-sectional studies of consumers' and business customers' switching behavior (e.g., Ganesh, Arnold, and Reynolds 2000; Keaveney 1995). For example, Heide and Weiss (1995) find that an organizational buyer's decision to switch to a new vendor of a high-technology product depends on his or her perception of rapid technological change, his or her prior experiences with vendors, and buying process formalization. However, we could find no studies that examine

---

Ruth N. Bolton is W.P. Carey Chair and Professor of Marketing, W.P. Carey School of Business, Arizona State University, and Honorary Professor, University of Groningen (e-mail: ruth.bolton@asu.edu). Katherine N. Lemon is Associate Professor of Marketing, Carroll School of Management, Boston College (e-mail: katherine.lemon@bc.edu). Peter C. Verhoef is Professor of Marketing, Faculty of Economics, Department of Marketing, University of Groningen (e-mail: p.c.verhoef@rug.nl). The authors are listed in alphabetical order and contributed equally to the manuscript. The Center for Customer Relationship Management at Duke University financially supported this research. The authors thank the anonymous cooperating company for supplying the data. They also thank Stijn van Osselaer for some theoretical suggestions. P. Rajan Varadarajan served as guest editor for this article.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (January 2008), 46–64

the antecedents of B2B customer upgrade behavior (for an overview of prior research in this area, see Table 1).

We address this gap in the literature by developing a model of how firms make upgrade decisions for fixed-price service contracts. We propose that the firm's decision to upgrade is influenced by the decision maker's perceptions of the supplier (i.e., at the relationship or account level), contract-level service experiences with the supplier, and interactions between supplier- and contract-level variables. Our study context is the repurchase of system support services by large firms. The firm's decision to upgrade a contract is represented by a binary logit model with random parameters for the contract-level variables and fixed parameters for the supplier-level variables. We estimate the model with cross-sectional, longitudinal data from a supplier of system support services with 120 business customers in Germany and the United Kingdom, representing a total of more than 2000 contracts.

Our study contributes to research on business customer–supplier relationships in several ways. First, it highlights the importance of understanding the relationship at both the supplier (i.e., account) level and the individual contract level because it demonstrates that both dimensions influence the firm's upgrade decision. Second, the results suggest a complex role for customer satisfaction, product–customer fit, and service quality, providing novel opportunities for the supplier to use sales and service functions to upgrade customers to higher (more expensive and potentially profitable) levels of services. Third, compared with prior research, the results show significant differences between the factors that influence the firm's decision to upgrade a contract and those that influence the firm's decision to renew a contract. Fourth, from a managerial perspective, the results suggest specific actions firms can take to influence customer upgrades. Finally, the data examined in this study are of the type available to many B2B firms, so it is possible for practitioners to apply our model to their individual business contexts.

Literature Review and Model Development

In this section, we develop a model of the firm's decision to upgrade a service contract conditional on the decision to renew the contract. In other words, the firm has already decided to repurchase a service contract from the same supplier; our model describes the firm's decision about whether to purchase a contract at the same service level or higher.[1] The factors affecting service contract renewal, service usage, and cross-buying (including upgrading and downgrading) are not always the same, and when the factors are the same, the predicted effect (positive, negative, or neutral) is not always the same (Bolton, Lemon, and Verhoef 2004). Thus, by framing the firm's decision problem in this way, our research isolates the factors that uniquely affect the upgrade decision. From the firm's perspective, the upgrade decision is the process of solving the problem of market matching and congruence (Alderson 1965). In essence, the firm and the supplier are engaging in dynamic adaptation (Dickson 1992); that is, the supplier has distinct levels of market offerings, and firms are heterogeneous with respect to their specific market needs. The firm engages in a problem-solving process to find the level of contract (appropriate level of service from the supplier) that best meets its needs. In this study, we examine specific factors the firm may use to enhance this market matching process.

All our predictions regarding the upgrade decision are made under ceteris paribus conditions—that is, after we control for other factors. In our discussion, we distinguish between factors associated with the specific contract and factors associated with the supplier. At the contract level, we propose that firms are influenced by service quality and price. At the supplier level, we propose that firms are influenced by decision-maker perceptions of the buyer–seller relationship—especially assessments of satisfaction and the criticality of the firm's needs—which may be considered account characteristics. We hypothesize interactions between firm- and contract-level variables. The model also includes covariates that control for relationship characteristics (share of customer, relationship duration) and contract type. By incorporating share of customer and price discount, we also account for competitive effects, which tend to be stronger when share is low or price discount is high.

Service Quality

Firms' prior service experiences, especially service quality, at the contract level are likely to influence the decision to upgrade a service contract (Zeithaml 1988). Unlike an initial purchase decision, the firm's upgrade decision is likely to depend on service quality under the current contract (see Kalwani and Narayandas 1995). The firm has many opportunities to assess the service during its experiences with suppliers, thereby learning to make more effective purchase decisions (Dodgson 1993). Organizational norms about the quality of service evolve through ongoing transactions and are based on prior experiences (Coleman 1990). Thus, when a focal contract has poor service quality, the firm is likely to conclude that the service level on the focal contract is inadequate for its needs, so it is less likely to renew the contract (Bolton and Myers 2003). Moreover, when competition exists and switching costs are low, suppliers that deliver consistently bad service are forced to exit the marketplace.

At the same time, unlike many other purchase decisions, the benefits of upgrading a service contract may seem greater when the firm has experienced poor service quality. There are three reasons poor service quality can result in upgrade. First, attribution theory suggests that people are likely to make causal attributions based on their prior service experiences (Folkes 1988). Thus, we believe that if other experiences are satisfactory and the firm intends to renew (i.e., ceteris paribus), the decision maker is likely to attribute poor service quality on the focal contract to an inadequate level of service contract. Second, prior research indicates that anticipation of regret may influence repeat purchase decisions such that anticipation of regret associ-

---

[1] Firms may also decide to "downgrade" or purchase a lower level of service; we find no incidence of downgrading in this context and therefore reserve the issues associated with downgrading for further research.

TABLE 1
Studies of Repeat Purchases and Service Quality

	Boulding et al. (1993)	Kalwani and Narayandas (1995)	Rust, Zahorik, and Keiningham (1995)	Rust et al. (1999)	Ganesh, Arnold, and Reynolds (2000)	Bolton and Myers (2003)	Li, Sun, and Wilcox (2005)	Ngobo (2005)	Bolton, Lemon, and Bramlett (2006)	Bolton, Lemon, and Verhoef (2006)
Topic	Model of service quality and repeat purchase intentions as function of "will" and "should" expectations	Investigation of whether relational approach to serving customers "pays off" in sales and profitability	An investigation of return on service quality	Model of customer expectation updating, focusing on distributions of expectations	A comparison of characteristics of consumers who switched banks and consumers who stayed loyal	Global market segmentation for services	Model of how customer demand for multiple products from a single provider evolves over time	An investigation of why some consumers migrate upward ("up-buy") and some migrate downward ("down-buy") in a service relationship over time	Service contract renewal as a function of average service levels, variability in service levels (especially extreme outcomes), and timing of service delivery	Firm's service contract upgrade decision modeled as a function of perceptions of its relationship with a supplier at the enterprise level, contract-level experiences, and interactions between firm- and contract-level variables
Analysis	Longitudinal experiment and cross-sectional survey	Cross-sectional and longitudinal	Cross-sectional	Cross-sectional experiment	Cross-sectional	Cross-sectional	Longitudinal	Longitudinal and cross-sectional	Longitudinal	Longitudinal
Dependent Variables	Repeat purchase intentions	Changes in net sales, return on investment, and other variables	Market share and net present value derived from consumer repeat purchase intentions	Perceived quality, intentions, word of mouth	No single dependent variable	Sales price elasticity for low-support contracts and sales price elasticity for high-support contracts	Buy/not buy banking products (e.g., loan, debit card, money market account)	Stay/defect, migration intentions (subscribe or not)	Renewal decision for a single contract	Upgrade decision for a single contract, conditional on renewal
Sample	Business-to-consumer (B2C): (1) hotel business trip, (2) educational institution	B2B, COMPUSTAT data for selected Standard Industrial Classification codes (n = 114)	Illustrative B2C example, hotel chain (n = 7882)	B2C, 200 undergraduate students shown scenarios describing restaurant service	B2C, banks (n = 200)	B2B: Canada, Germany, Japan, Korea, Singapore, United Kingdom, United States (508 low support, 445 high support)	B2C, banking panel data (n = 1201)	B2C, theatergoers (n = 200)	B2B: Germany and United Kingdom (2442 contracts)	B2B. Germany and United Kingdom (2076 contracts)[a]
Unit of Analysis	Individual consumer level	Comparison of matched samples over time	Individual-level model linked to aggregate model	Individual consumer level	Comparisons of groups	Enterprise level	Individual household level	Individual consumer level	Contract level	Contract level

TABLE 1
Continued

	Boulding et al. (1993)	Kalwani and Narayandas (1995)	Rust, Zahorik, and Keiningham (1995)	Rust et al. (1999)	Ganesh, Arnold, and Reynolds (2000)	Bolton and Myers (2003)	Li, Sun, and Wilcox (2005)	Ngobo (2005)	Bolton, Lemon, and Bramlett (2006)	Bolton, Lemon, and Verhoef (2006)
Nature of Predictor Variables	(1) Manipulation or perception of service performance attributes and (2) 36 items from SERVQUAL	Dichotomous categorization: relational versus transactional approach to service	Perceptual ratings of service quality attributes	Service variability (participants were shown graphs of up to 20 past experiences)	Perceptual ratings of service quality attributes	Service operations measures, organizational characteristics, national/cultural variables budget	Accounts with competitors, overall satisfaction, switching costs, demographics	Perceptual ratings of overall satisfaction, service quality and pricing policy, relationship variables, demographics	Service operations measures, price, discount	Decision-maker perceptions (satisfaction, criticality), service operations measures, price, discount, share of customer, relationship duration
Service Quality Constructs	(1) Overall quality of hotel (unfavorable/favorable) and (2) SERVQUAL dimensions of reliability and empathy	Single dimensions: relational or transactional approach to service	Grouped by business processes and their underlying subprocesses	Overall service quality (unfavorable/ favorable) following Boulding and colleagues (1993)	Three service factors: people, convenience, ease of transactions (people factor is different across groups)	Responsiveness, reliability, and assurance (each represented by two measures)	None specifically: perceptions of overall satisfaction with the bank	Principal components analysis yielded three factors: quality of acting and decor, quality of performance, quality of subscription system	Design quality (contract type) and experience quality (engineer work minutes)	Overall quality (represented by mean resolution time per contract)

---

[a] This sample is smaller than the sample Bolton, Lemon, and Bramlett (2006) use because (1) satisfaction measures were not available for some firms and (2) high-level contracts are not eligible for upgrade

ated with dropping a service can influence a customer to retain the service (Inman and Zeelenberg 2002; Lemon, White, and Winer 2002). Extrapolating to upgrade decisions, we believe that when a decision maker experiences poor service quality, anticipated regret is likely to cause forgone alternatives, such as a higher level of service contract, to become more attractive. Third, Park, Jun, and MacInnis's (2000) experimental research shows that when customers are committed to buying within a category, "subtractive framing" (in which customers are asked to delete options from a fully loaded model) encourages them to choose more options. In our study context, poor service quality on the focal contract may encourage subtractive framing and, thus, upgrading. This is also consistent with the negativity effect, which suggests that a single negative piece of information is often weighed more heavily than a single positive piece of information in consumer evaluations (Price 1996). Finally, this is also consistent with learning theory, which suggests that firms (Dodgson 1993) and consumers (Janiszewski and Van Osselaer 2000) engage in trial-and-error behavior over time that results in learning. This implies that, over time, firms may eventually learn the appropriate contract match for their needs.

In summary, prior research suggests that a firm's service quality has a positive relationship to customer retention; that is, compared with other contracts the firm holds, a firm is less likely to renew a service contract when there is poor service quality on the focal contract. However, the preceding discussion suggests that there are theoretical reasons for this, as follows:

> H_1: Conditional on renewal, a firm is more likely to upgrade a service contract when the quality of service on the focal contract is poor.

As we discuss subsequently, we expect that certain factors will moderate the effect of service quality on the firm's upgrade decision.

Degree of Product–Customer Fit: Criticality

Studies of B2B purchasing behavior have long distinguished between products that have a large impact on the firm's financial performance and products that have a small impact on this performance (Kraljic 1983). For example, Cisco Systems does not outsource "mission-critical" activities, defined as "activities that, if performed poorly, would pose an immediate risk to the company" (Bridge 2004). Researchers also recognize that business activities that are "critical" or important to a firm's profit streams are less likely to be outsourced (Stremersch et al. 2003). High-level service contracts are designed for critical service environments, whereas low-level service contracts are designed for less critical service environments. Consequently, we predict that firms with a decision maker who perceives a service as critical will be more likely to upgrade from a low-level service contract. For example, retailers to which copiers may be critical to their service operations (e.g., FedEx Kinko's) are more likely to upgrade their copier maintenance contracts from a low to a higher level of service than other retailers with less critical needs.

The underlying theoretical rationale for our prediction is that a poor fit between the product and the customer should positively influence the firm's decision to upgrade. This is consistent with the idea of market matching (Alderson 1965; Dickson 1992), suggesting that the firm will adapt and upgrade the contract to create a better fit with its needs. Following Kristof (1996, pp. 4–5), we define product–customer fit as the degree of compatibility between customer and product that occurs when the product provides what the customer needs.[2] Product–customer fit is poor when a customer perceives service as critical (i.e., a high level of service is required) but purchases a low-level service contract. Mittal and Kamakura (2001) discuss the nature of the relationship (or fit) between the customer and the brand, finding that customers with different characteristics have different satisfaction thresholds and, therefore, different probabilities of repurchase. Poor fit will decrease the likelihood that a contract is renewed because the service does not appear to deliver on its promise, but at the same time, it will increase the likelihood that a contract is upgraded because the firm recognizes the poor fit and adjusts the contract provisions.

> H_2: Conditional on renewal, a firm's likelihood of upgrading a service contract is positively related to the decision maker's perception of the criticality of the service to the firm's operations.

Moderating Effect of Criticality

Recent research indicates that when people pay more attention to a product, service quality has a larger effect (more positive) on loyalty (Bloemer and Kasper 1995). This finding suggests that decision makers who perceive critical needs weigh contract-level service experiences more heavily in making their upgrade decision, which results in a "hyperfocus" on these contract decisions. For this reason, when the service is perceived as critical, firms are more likely to upgrade when the focal contract service quality is low. In contrast, if the service is not perceived as critical, firms are less likely to upgrade in response to poor service quality at the contract level (because decision makers are paying less attention to all contracts).

> H_3: Conditional on renewal, the strength of the relationship between a firm's likelihood of upgrading a service contract and service quality on the focal contract is greater when the decision maker perceives the service as critical than when he or she perceives it as noncritical.

Satisfaction with Prior Service

Customer satisfaction with prior service experiences has a positive effect on the length, breadth, and depth of the customer–firm relationship (e.g., Oliver 1997). In a field investigation of B2B relationships in mature industrial markets, Narayandas and Rangan (2004) report that a decision maker's favorable evaluations of performance within the contract terms (as well as outside them) lead the firm to increase its commitment to the supplier. Because the upgrade decision is an increase in the firm's commitment to the existing relationship, we predict that a firm with a decision maker who is highly satisfied with the supplier (i.e., he or she has high cumulative satisfaction across all aspects of the relationship) will be more likely to upgrade.

---

[2] See also Vancouver and Schmitt (1991) regarding person–organization fit, as well as the strategy literature, in which product–market fit is also sometimes called "congruency."

> H_4: Conditional on renewal, a firm's likelihood of upgrading a service contract is positively related to the decision maker's overall satisfaction with the existing supplier.

Our hypothesis is consistent with the well-established relationship between satisfaction and repeat purchase behavior. However, it is also possible that firms that are satisfied with the supplier are less likely to upgrade because they believe that the current service level is exactly meeting their needs.

Moderating Effect of Satisfaction: The Rose-Tinted-Glasses Effect

Research in psychology and consumer behavior suggests that people exhibit a confirmatory bias (Hoch and Ha 1986), such that their prior opinions or expectations influence not only their overall opinions but also their perceptions of new information (Oliver and Burke 1999). Boulding, Kalra, and Staelin (1999) find that confirmatory bias decreases with higher levels of experience and lower levels of product complexity and ambiguity. Nayakankuppam and Mishra (2005) find that sellers (people who currently own a product) overrepresent the positive features of the product compared with buyers—a "rose-tinted-glasses" effect (Hendrick and Hendrick 1988). Because we are studying complex B2B services, these findings suggest that the firm's assessment of service quality will be positively biased for satisfied decision makers and negatively biased for those who are dissatisfied. Thus, we believe that firms will weigh contract-level service quality less heavily in their upgrade decision when decision makers' satisfaction levels are high.

> H_5: Conditional on renewal, the strength of the relationship between a firm's likelihood of upgrading a service contract and service quality on the focal contract is weaker when the decision maker's overall satisfaction with the supplier is high than when it is low.

Price Effects

When customers consider whether to upgrade a service contract, the price of the service may also influence their decision making. Prior research suggests that a customer's reference price will influence his or her decision to purchase a good or service. The reference price is the price against which buyers compare the offered price of a product (Winer 1986). The concept of dual entitlement (Kahneman, Knetsch, and Thaler 1986) suggests that consumers believe that firms are entitled to a reference profit, that customers are entitled to a reference price, and that price increases commensurate with cost increases will be perceived as fair (e.g., Bolton and Alba 2006). Similarly, in this research context, an upgrade represents more benefits to the firm and higher costs to the supplier. Therefore, firms that upgrade to a higher level of contract most likely pay a higher price for that contract than if they had not upgraded. Thus, for completeness, we predict the following:

> H_6: Conditional on renewal, a firm's likelihood of upgrading a service contract is positively related to the current price of the service contract.

Bolton and Myers (2003) find that customers who receive less reliable service over time are more price sensitive than customers who receive more reliable service. This suggests that price should moderate the effect of service quality on the firm's upgrade decision.

> H_7: Conditional on renewal, the strength of the relationship between a firm's likelihood of upgrading a service contact and service quality on the focal contract is weaker when the price of the contract is high than when it is low.

Summary

We hypothesized that service quality on the focal contract (SQual, H_1), perceived criticality (Criticality, H_2), satisfaction (Satisfaction, H_4), and contract price (LogPrice, H_6) influence the decision to upgrade a service contract. We also hypothesized that perceived criticality, customer satisfaction, and contract price moderate the effect of service quality on the customer's decision to upgrade (H_3, H_5, and H_7). Prior research also suggests that models of ongoing buyer–seller relationships should account for certain covariates. Thus, we include supplier-level variables to capture relationship characteristics, such as commitment (which can be represented by measures of share of customer [Share] and customer relationship duration [Duration]), contract-level variables for price discount (LogDiscount) to account for linear effects that have been previously established (Verhoef, Frances, and Hoekstra 2001), and a dummy variable for contract type (ContractType). Figure 1 depicts the conceptual model of the upgrade decision.

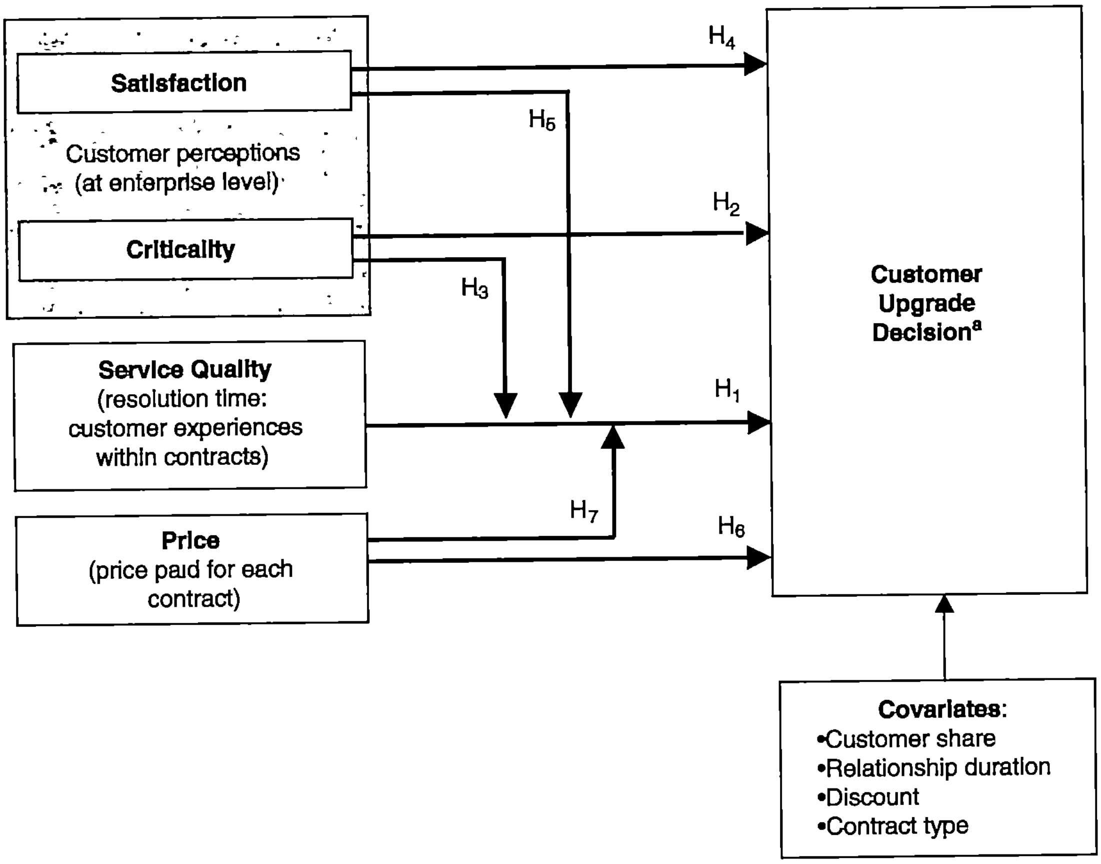
Overall, the firm's decision to upgrade a contract can be considered in terms of the overall assessment of value associated with the contract—that is, how the firm weighs the benefits and costs of the upgraded contract. A customer's assessment of value from a service has been defined as the customer perceived trade-off between costs and benefits (Brady et al. 2005). In a B2B context, Bolton and Drew (1991) suggest a cost–benefit framework in which a customer's assessment of value depends on benefits received and sacrifice (i.e., monetary and nonmonetary costs associated with using the service; e.g., Zeithaml 1988) and the customer's frame of reference or context (Holbrook 1999). At a high level, our model is consistent with the view that firms weigh the benefits (service quality, satisfaction) and costs (price) associated with the contract when determining whether to upgrade and account for their own frame of reference (criticality, share of requirements, relationship duration, and contract type).

Model Specification and Estimation Procedure

Following a long tradition of theoretical and empirical work modeling firm behavior (e.g., Coughlan 1985; Heide and Weiss 1995), we represent the firm's decision to upgrade a service contract by a binary choice (logit) model. The firm (i) renews a service contract (k) by choosing the alternative (upgrade or not) with the highest expected future value (V_{ik}). That is,

$$\text{Prob}(\text{Upgrade}_{ik}) = \text{Prob}(V_{ik^*} > V_{ik}). \quad (1)$$

FIGURE 1
Conceptual Model of Customer Upgrade Decision


Satisfaction
Customer perceptions (at enterprise level)
Criticality
Service Quality (resolution time: customer experiences within contracts)
Price (price paid for each contract)
H4
H5
H2
H3
H1
H7
H6
Customer Upgrade Decision^a
Covariates:
•Customer share
•Relationship duration
•Discount
•Contract type

^aControlling for the renewal decision.

In this situation, a firm makes repeated decisions (across contracts), creating dependent observations. The number of contracts purchased typically differs between companies, so that we expect the data to be "unbalanced." To account for the interdependency within each firm and to allow varying coefficients across firms, we estimate a random parameters model (Train 2003). Specifically, we estimate random coefficients for the contract-level variables (Z_{ik}) and fixed coefficients for the supplier-level variables (X_i). This specification is known as a mixed-effects logit model.

If a firm i makes k contract upgrade decisions, we formulate our model as follows:

$$V_{ik} = \alpha' X_i + \mu'_{ik} Z_{ik} + \varepsilon_{ik}, \tag{2}$$

where α' represents a vector of fixed parameters for the supplier-level data X_i, μ'_{ik} is a vector of random coefficients for the contract-level data Z_{ik}, and ε_{ik} is i.i.d. extreme value. The terms in Z_{ik} are error components that, along with ε_{ik}, define the stochastic portion of the utility (Train 2003). As such, we define the unobserved random portion of utility (η_{ik}) as

$$\eta_{ik} = \mu'_{ik} Z_{ik} + \varepsilon_{ik}. \tag{3}$$

We assume that the mean and standard deviations of the random parameters are normally distributed.

In summary, we can write our mixed-effects binary logit model as follows:

$$\begin{aligned} \text{Upgrade}_{ik} = f(&\text{Satisfaction}_i, \text{Criticality}_i, \text{SQual}_{ik}, \\ &\text{SQual}_{ik} \times \text{Satisfaction}_i, \text{SQual}_{ik} \times \text{Criticality}_i, \\ &\text{LogPrice}_{ik}, \text{SQual}_{ik} \times \text{LogPrice}_{ik}, \\ &\text{LogDiscount}_{ik}, \text{Share}_i, \\ &\text{Duration}_i, \text{ContractType}_{ik}) + \varepsilon_{ik}. \end{aligned} \tag{4}$$

Recall that we estimate fixed coefficients for the supplier-level variables (X_i) and random coefficients for the contract-level variables (Z_{ik}). The interactions between perceptions at the supplier level (Satisfaction_i, Criticality_i) and service quality at the contract level (SQual_{ik}) and the interaction between price (LogPrice_{ik}) and service quality create contract-level variables, so we estimate random parameters for these interaction terms.

We specify and estimate our model conditional on contract renewal; that is, we estimate the model only for the 2076 contracts that were renewed. However, the (potential) hierarchical structure of the renewal and upgrade decisions might suggest the specification of a nested logit model that allows for a correlation between the error terms of the unobserved utilities. Because we do not use a nested logit model, it might be argued that we are biasing the estimates of our logit model parameters (Ben-Akiva and Lerman 1985). However, this argument ignores the issue of random parameters arising from unobserved customer (i.e., firm) heterogeneity. It is cumbersome to account for customer heterogeneity in a nested logit when both upgrade and renewal decisions depend on (mostly) the same explanatory variables. More important, Louviere (2003) argues that if the researcher uses certain mixed-effects logit models (e.g., with a hierarchical structure, random effects), it may not be possible to use statistical tests to identify the "correct" model specification. On the basis of these considerations, we chose a reasonably parsimonious, binary logit model with mixed effects and estimated it with data describing renewed contracts only.

Method

Study Context

The study context is computing system support services purchased by large European firms. Firms spend (approximately) €650,000 per year on computing system support services. The service contracts are fixed-price contracts for one year; customers are not billed on service-usage levels. Firms purchase system support contracts independently for systems that have been purchased previously, and an individual contract is uniquely associated with a particular system. Most firms purchase from multiple suppliers. There are eight suppliers in the European market, and firms may also use internal support services. Large firms usually have many systems, and they buy multiple contracts from multiple suppliers.

Suppliers in many industries give the different levels of their service contracts descriptive names, such as "mission-sensitive support," "proactive support," or "managed services," that are intended to help the buyer achieve a good product–customer fit. System support is provided out of centralized facilities, so that it is standardized across customers and contracts. In our study, the supplier offers contracts that are available at three levels of system support, which we call low, medium, and high support. We do not study high support (which cannot be upgraded); we study upgrades from low to medium levels or from medium to high levels. For each level of support, the supplier has specific, contractually defined service performance levels, such as "24 hours, 7 days a week availability, with a guaranteed response within 2 hours." Each level of support provides a subset of the services of the next (higher) support level. A higher level of support contract is more expensive than a lower level of support contract, but it provides higher levels of support that can reduce resolution times, thus increasing service quality. Importantly, over time, a customer may decide to upgrade the support level of a contract. No customers downgraded their contracts in our database.

The Database

For our study, a cooperating supplier provided a probability sample of customer records that described purchases of system support service contracts by 120 large firms (i.e., business customers) in Germany and the United Kingdom. The sampling procedure ensured that the sampled firms operated the same system, thus reducing the effects of variables extraneous to our study (e.g., factors unique to after-sales service for a particular system).[3] The data described the firm–supplier relationship over a three-year period (1997–1999). We combined information from four primary sources: (1) a master file describing the characteristics of each customer account, including number and type of contracts; (2) annual billing data for all contracts of each customer, including price and any discounts; (3) internal monthly service operations data (i.e., service quality metrics) for all contracts of each customer; and (4) self-report data regarding satisfaction and criticality from a survey of the decision maker at each firm. (The survey was administered to the person identified by the supplier as the decision maker, after he or she was rigorously screened by telephone to ensure that he or she either recommended or made the final decision on system support contracts.[4]) We merged these sources to create a comprehensive database that incorporates both cross-sectional and longitudinal observations at the customer account and contract level. Note that we measure the predictor variables (Satisfaction_i, Criticality_i, SQual_{ik}, Share_i, Duration_i) in the period before the upgrade decision, so it is possible to make causal inferences from our data. The service contract's price and any discount are from the time of purchase.

All firms purchased five or more contracts from the same supplier; the average number of contracts was 19.7, and the standard deviation was 23.7. Of the 120 firms, 43.3% upgraded at least one contract—upgrading occurs only if firms renew contracts (there were no firms in our database that downgraded contracts); of the contracts, 88% (2076) were subsequently renewed, and 8% (165) of the renewed contracts were upgraded. Thus, our analysis database contains the 2076 renewed contracts for the 120 firms. The relatively low percentage of upgraded contracts is likely because most firms purchase a service level that is appropriate for their needs. We conducted preliminary statistical analyses and determined that the upgrade process was the same for both contract levels (low to medium or medium to high); overall, more medium-level contracts were upgraded (18%) than low-level contracts (5%). There was no evidence of a need for interaction effects between contract type and any of the other predictor variables. Thus, we capture this feature of the data by incorporating a dummy variable that represents medium-level contracts. We discuss this issue further in the "Results" section.

---

[3] The cooperating company drew the sample with the restriction that the equipment (covered by the service contract) was still in use and would continue in use, thus removing any cases in which the customer no longer owns or uses the equipment and thus does not renew because the equipment is no longer in service.

[4] Service operations data (i.e., resolution time measures) were collected from January 1997 to February 1999; billing records (i.e., purchase data, including upgrade decision [dependent variable], contract type, and price information) were collected in February 1998 and January 1999; the survey data (i.e., satisfaction, criticality, duration, and customer share measures) were collected in July 1998.

Measures

Table 2 reports the predictor variables, their measures, the data source and level (i.e., entire customer account or contract), and descriptive statistics (average, standard deviation, and skewness). The average firm had been purchasing from the supplier for six years and spent 30% of its total system support budget on the supplier's service contracts. Because the majority of the predictor variables were measured in a straightforward way, we discuss only measures that require additional explanation.

Satisfaction. A survey elicited self-report measures from chief information officers or management information system managers who were identified from company records and screened by telephone to ensure that they either recommended or made the final decision on system support contracts. We calculated the average of three survey items to measure the decision maker's satisfaction. The items are similar to items used in other studies (e.g., Bolton 1998; Oliver 1997). They elicit perceptions of (1) the decision maker's overall satisfaction with his or her relationship with the supplier, (2) the value that the supplier adds to the firm's business activities, and (3) the overall satisfaction with customer service. We assessed each question on a five-point scale (1 = "extremely unsatisfied," and 5 = "extremely satisfied").

Criticality. We measure the criticality of the service to the firm by asking the decision maker, "When your system fails, what effect does that have on your business ("not at all serious/very serious")?" Because the distribution of responses was heavily skewed, we collapsed the multiple-response variable to a dichotomous variable based (approximately) on a median split. Thus, the variable is coded as 1 if the service is highly critical and as 0 if otherwise. In total, 71% of the firms consider their services critical for their operations.

Contract-level service quality. Before this study, the supplier had hired a market research company to conduct research with its European business customers. The company conducted face-to-face interviews with 40 decision makers and administered large-scale telephone surveys to decision makers with the dual goals of understanding perceptions of service quality and areas for improvement. The supplier also participated in an industrywide benchmarking study that showed that its satisfaction ratings were roughly comparable to the other two major suppliers in Europe. However, timeliness, completeness, and effectiveness of system support consistently emerged as frequent customer concerns. Specifically, respondents identified speedy resolution time as a key aspect of service quality that they considered in repeat purchase decisions. The following are verbatim comments from four decision makers:

> [The supplier needs] to have a more timely response; it would be better to get rid of the callback system—replacing it with a direct connect [to an engineer].
>
> I have a highly technical staff and [the supplier] should recognize that ... we go through our own diagnosis and exhaust our resources. We get a junior helper [at the customer support center] and have to go through the same process with [the supplier] and waste our time.
>
> [The supplier] should shorten the resolution time.
>
> [The supplier] needs to keep up with all the information that we provide them on our problems and resolve them in a more timely manner.

These decision makers' comments reflect the notion that additional supplier resources (e.g., faster connect time, more customer support staff, access to knowledgeable engineers) were necessary to reduce resolution time, thus reducing system downtime and its negative effect on business performance.[5] Preliminary statistical analyses also supported the notion of a single variable—resolution time—as an indicator of service quality. Respondents' intense dissatisfaction with a single service quality attribute is consistent with recent research. Slotegraaf and Inman (2004) report that as customers approach the end of a service contract period, satisfaction with attributes that can be remedied declines at a faster rate. In our study context, concerns about slow resolution time can be remedied by upgrading the service contract so that resolution time becomes a prominent dissatisfier and a key input into the upgrade decision.

For these reasons, we measured service quality at the contract level by calculating the average resolution time per service request for the focal contract (expressed as minutes/request) from the monthly service operations data during the year before the purchase decision.[6] A large value of this variable indicates poor service quality (rather than good service quality). We calculated the log of this measure to create a variable with a less skewed distribution. Thus, our model incorporates service quality measured as the log of average resolution time per request for the focal contract. This transformation implies that as resolution time becomes larger, there are diminishing marginal effects of service quality at the contract level.

Price. We obtained the price of each contract from the billing data and measured it at the contract level. We included two aspects of price in the model: contract price (in dollars) and price discount (expressed as a percentage)

---

[5] Note that resolution time is not the same as response time. Resolution time is the time elapsed until the customer request has been resolved, not the time until first response. However, not surprisingly, quantitative analyses of the operations data (not shown herein) indicated that response time, engineer work time, number of repeat requests, and other service quality/business process metrics were correlated with resolution time (and they might be considered its antecedents).

[6] Our approach is a different from cross-sectional studies that rely on perceptual measures of service quality (which can be standardized across companies and industries). Because we study firms that purchase from a single supplier, we can measure service quality using business process metrics from the service operations database. Our approach is consistent with prior within-company research that uses business process metrics to represent service quality, thus producing actionable results for managers (e.g., Bolton and Myers 2003). In models estimated with perceptual measures, it is frequently possible to capture multiple dimensions of service quality, such as reliability, responsiveness, assurance, empathy, and tangibles. However, our statistical tests indicated that a single measure of service quality was sufficient to explain the upgrade decision. The inclusion of additional variables did not improve the explanatory power of the model.

TABLE 2
Measures and Descriptive Statistics (N = 2076)

Variable (Name)	Description and Measurement[a]	Data Source	Data Level	Average[b]	Skewness
Upgrading (Upgrade_{ik})	Contract upgraded or not. A dichotomous variable determined by comparing billing records for January 1999 and January 1998	Billing	Contract	.079 (.27)	3.11
Satisfaction[c] (Satisfaction_i)	Average of the following questions (Coefficient $\alpha = .70$) How satisfied … (1 = "extremely unsatisfied," 5 = "extremely satisfied") •as a whole with the customer service which you buy from [the supplier]? •with the value that [the supplier] adds to the business activities? •are you overall with your relationship to [the supplier]?	Survey	Supplier	4.04 (.55)	–.86
Service quality (SQual_{ik})	Log of the average resolution time (in minutes) for the focal contract: 1998 average resolution time per request for focal contract	Operations	Contract	.38 (1.09)	3.11
Criticality[c] (Criticality_i)	"When your system fails, what effect does that have on your business?" Dichotomous variable derived from self-report data (1 = "very serious")	Survey	Supplier	.71 (.46)	–.90
Covariates					
Share of customer (Share_i)	Percentage of service contracts purchased by customer from the supplier. Calculated as: the total dollar value of contracts purchased from the supplier (from the billing data) divided by the dollar value of the total support budget (as reported by the decision maker)	Survey, master file	Supplier	.30 (.22)	1.07
Relationship duration (Duration_i)	Duration of relationship in years. Calculated as the years between January 1999 and start of relationship	Master file	Supplier	6.41 (3.19)	.15
Price (LogPrice_{ik})	Log of the price level per contract. Dollars per contract (scaled for reasons of confidentiality)	Billing	Contract	.74 (1.55)	–.56
Discount (LogDiscount_{ik})	Log of the price discount per contract. Discount is expressed as a percentage of the dollar price per contract	Billing	Contract	–4.46 (4.04)	–.32
ContractType_{ik}	Contract level (medium = 1, otherwise = 0)	Master file	Contract	.16 (.36)	1.88

[a] "Supplier" should be interpreted to mean the business customer's perception of the supplier or service.
[b] Standard deviation is in parentheses.
[c] From a survey administered on July 1998.

for each contract. To allow for nonlinear effects of price and discount while minimizing the potential for multicollinearity, we used a log form for the price variables. This formulation is more parsimonious than including both linear and quadratic price variables.

Covariates. We measured two covariates at the customer account level: share of customer and relationship duration. We calculated share of customer by dividing the dollar value of the number of contracts purchased from the supplier by the dollar value of the firm's total system support service budget (as reported in the survey). The firm–supplier relationship duration is a self-report measure from the survey; it was measured in years. (Note that it is not the same as the length of the decision maker's relationship with the supplier.) Finally, we incorporated a dummy variable to represent contract type (1 = medium, and 0 = otherwise).

Table 3 displays the correlation matrix of the predictor variables. The majority of the coefficients are low. An exception is the correlation (.30) between list price and percentage price discount, suggesting that contracts with higher list prices have higher percentage discounts. This correlation is not surprising given that the European market for system support contracts is highly competitive. Overall, we concluded that multicollinearity among the predictor variables is reasonably low.

Model Estimation

We estimated our model using LIMDEP 8.0, which provides accessible estimation procedures for the random parameter versions of the binary logit model (Greene 2003). We performed simulation in the random parameter model using 100 Halton draws (Greene 2003). We estimated three models: (1) a covariates-only model, (2) a main effects–only model, and (3) a full model including the interaction effects. The fixed coefficients, random coefficients, scale parameters of the random coefficients, and fit statistics of these three models appear in Table 4. This table shows significant scale parameters for the random coefficients for multiple variables. The significant scale parameters imply that there is heterogeneity in the estimated parameters, so we conclude that accounting for heterogeneity with random coefficients is a useful modeling strategy. This feature is important to remember when interpreting the estimated parameters. For example, the scale parameter for LogPrice is rather large, suggesting significant heterogeneity across the sampled firms. Thus, for some customers, the price effect may be smaller, whereas for other customers, the price effect may be larger.

Results

Model Fit

Of the three estimated models, Models 2 and 3 have a much better fit than Model 1. Model 3 has a higher likelihood and R-square than Model 2 and an equal Akaike information criterion (AIC). However, note that AIC imposes a heavy penalty in a random coefficients model because the addition of each predictor variable adds more than one parameter estimate.

We believe that the full model (Model 3) is preferable for three reasons. First, most of the hypothesized predictor variables in the upgrade model are significant ($p < .05$). Second, according to log-likelihood ratio test, the estimated model is highly significant ($p < .0001$). Third, for comparison purposes, we estimated a model that excluded the contract-level variables with their random coefficients. That is,

$$(5) \qquad \text{Prob}(\text{Upgrade}_{ik}) = g(\text{Satisfaction}_i, \text{Criticality}_i, \text{Share}_i, \text{Duration}_i) + \varepsilon_{ik}.$$

We compared our proposed model (Equation 4) with the "firm-only" model (Equation 5) and found that the proposed model produces an improved fit (AIC = 1130.28 versus 970).[7] Thus, our model comparisons suggest that a binary logit model with random coefficients, plus both firm- and contract-level variables, is an appropriate modeling strategy.

We assessed the stability of our estimated model (Equation 4) as follows: We estimated our model using 75% of our sample. We repeated this procedure three times (i.e., three random samples). Of the estimated parameters, 97.2% have the same sign as in our reported model. The significance levels are the same for 86% of the estimated parameters. Consequently, we conclude that our model results are stable.

Hypothesis Tests

We find strong support for the model. The results for the three hypotheses regarding main effects are as follows: Firms experiencing poor service quality at the contract level are indeed more likely to upgrade (H_1, $p = .013$). However, H_2 is not supported; the decision maker's perception of the criticality of the firm's service does not influence upgrading (H_2, $p = .24$). Firms with decision makers who are more satisfied with the overall relationship with the firm are more likely to upgrade contracts, in support of H_4 (H_4, $p < .001$). Current contract price has a positive effect (with diminishing marginal returns) on upgrading (H_6, $p < .001$). We predicted an interaction effect between service quality and criticality (H_3). When decision makers perceive the firm's service needs as critical, we expected that poor service quality at the contract level would lead to a higher probability of upgrading. This hypothesis was not supported (H_3, $p = .49$).

We predicted that there would be an interaction effect between satisfaction and service quality. This hypothesis was supported (H_5, $p < .001$). We conducted simulations for different satisfaction/service quality levels and discovered that this interaction effect captures the following feature: For poor service quality, business customers who have low

---

[7] We tested models with other functional forms, such as a nonlinear effect of satisfaction (i.e., quadratic). The fit of this model was worse than our model (AIC = 972 versus 970). We also considered whether firms make comparisons across contracts by including the difference between resolution time of the focal contract and the average resolution of all other contracts. The model fit of our reported model was much better than the model fit of the model accounting for a comparison across contracts (AIC = 996 versus 970).

TABLE 3
Correlation Matrix (N = 2361)

Variable	Upgrade$_{ik}$	Satisfaction$_{i}$	Criticality$_{i}$	SQual$_{ik}$	LogPrice$_{ik}$	LogDiscount$_{ik}$	Duration$_{i}$	Share$_{i}$	ContractType$_{ik}$
Upgrade$_{ik}$	1.00								
Satisfaction$_{i}$	.04	1.00							
Criticality$_{i}$	.05	.22	1.00						
SQual$_{ik}$	–.01	–.03	–.02	1.00					
LogPrice$_{ik}$	09	.16	.16	.01	1.00				
LogDiscount$_{ik}$	–.15	.06	–.14	–.13	.21	1.00			
Duration$_{i}$	–.04	.28	.15	–.03	.08	–.03	1.00		
Share$_{i}$	.07	.01	–.04	.09	.30	.09	.14	1.00	
ContractType$_{ik}$	.20	.03	.02	.34	.08	–.12	.02	.04	1.00

TABLE 4
Estimation Results (N = 2076)

Variable (Hypothesis)	Hypothesis	Expected Sign	Covariates-Only Model	Main Effects–Only Model	Full Model
Constant			–2.259**	–5.069**	–6.141**
Supplier Level					
Satisfaction$_i$	4	+		.498**	.712**
Criticality$_i$	2	+		–.061	–.159
Contract Level					
SQual$_{ik}$ (resolution time)	1	+		–.490**	1.179*
LogPrice$_{ik}$	6	+		.297**	.429**
LogDiscount$_{ik}$				–.131**	–.138**
Criticality$_i$ × SQual$_{ik}$	3	+			–.103
Satisfaction$_i$ × SQual$_{ik}$	5	–			–.592**
LogPrice$_{ik}$ × SQual$_{ik}$	7	–			–.241**
Covariates					
Share$_i$			.979**	.869**	.999**
Duration$_i$			–.062**	–.070**	–.159**
ContractType$_{ik}$			.799**	1.084**	1.011**
Scale Parameters					
Constant			1.097**	.305**	.126*
ContractType$_{ik}$			.943**	.631**	.964**
SQual$_{ik}$ (resolution time)				.507**	.429**
LogPrice$_{ik}$				.420**	.469**
LogDiscount$_{ik}$				.082**	.140**
Criticality$_i$ × SQual$_{ik}$					.720**
Satisfaction$_i$ × SQual$_{ik}$					.437**
LogPrice$_{ik}$ × SQual$_{ik}$					.217**
Model Fit					
Log-likelihood			–502	–471	–465
AIC			1016	970	970
McFadden R^2			.13	.18	.19

$^*p < 05.$
$^{**}p < .01.$
Notes: Two-sided *p*-values for hypothesis tests. Number of observations = 2076.

satisfaction are more likely to upgrade than those who have high satisfaction. However, for good service quality, customers who have high satisfaction are much more likely to upgrade than those who have low satisfaction. This finding is consistent with our hypotheses. Overall, high satisfaction leads to upgrading, except when there is poor service quality on the focal contract. We also predicted an interaction between current contract price and service quality (H_7). We find strong support for this effect, suggesting that poor service quality reduces the effect of price on the upgrade decision (negative interaction effect, H_7, $p < .001$).

Covariates

Firms that devote a larger portion of their total support budget to the supplier (i.e., higher "share of customer") are more likely to upgrade a contract ($p < .001$), and firms that have lengthy relationships with the supplier are less likely to upgrade ($p < .001$), perhaps because they have already achieved an optimal product–customer fit. In addition to the price effects we reported, percentage price discount has a negative influence on the upgrade decision ($p < .001$), with diminishing marginal returns. The significant effects of price in this study are rather surprising because price is often not a statistically significant predictor in models of customer retention (e.g., Bolton 1998; Bolton, Lemon, and Bramlett 2006). Contracts at the low service level are more likely to be upgraded ($p < .001$) than service contracts at the medium service level, ceteris paribus. This result raises the following question: Are the model results different when we omit the low-service-level contracts from the analysis? However, we conducted this analysis and still found that the effects are much the same as we reported previously.[8]

---

[8] Another approach would be to specify one model (Model A) to represent the decision to upgrade contracts from low to medium and another model (Model B) to represent the decision to upgrade from medium to high. However, there are some statistical reasons we did not choose to use this approach. First, the number of low-level contracts is relatively small compared with the number of medium-level contracts (325 versus 1751). Because only a small percentage of contracts are upgraded (8%), the number of low-

Theoretical Implications

Our results suggest that the customer upgrade decision is significantly different from the new product/service purchase decision, because prior experience with the customer is critical, and different from the customer renewal decision, because the firm's experience with the supplier appears to influence these decisions in different ways. Overall, we find that the results are consistent with the basic ideas of market matching and congruence. Firms use their experiences with the service contract to determine whether the current service contract fits their needs, and they upgrade when necessary to find the appropriate level of service they require from the supplier. The ways firms use this information to obtain the best match with their needs provide new insights into the firm's upgrade decision and our understanding of B2B customer behavior.

Accounting for Multiple Inputs into the Upgrade Decision

Traditional approaches to choice modeling focus on connecting information on the characteristics of an offering and prior customer experience to behavior (e.g., Guadagni and Little 1983). More recent attention has shifted to understanding how customers' overall attitudes toward and evaluations of an offering explain behavior (Bolton 1998; Gustafsson, Johnson, and Roos 2005; Mittal and Kamakura 2001; Verhoef 2003). Our study underscores the need to combine both approaches to understand the complexities of buyer behavior.

In particular, the results highlight the circumstances under which buyers may pay more attention to objective information (e.g., service quality, price) or summary evaluations (e.g., satisfaction) and the extent to which summary evaluations may actually interact with contract-level experiences to influence buyer behavior. Thus, at the macro level, the results of this study suggest that (1) both contract-level experiences and firm-level decision-maker perceptions influence contract-level decisions directly and (2) decision-maker perceptions of the supplier interact with contract-level experiences to influence these decisions. A model of firm decision making estimated solely at the supplier level or solely at the contract level omits interesting antecedents of the upgrade decision. This finding is particularly novel and important because the majority of prior B2B research has been conducted solely at the supplier or customer account level (i.e., eliciting measures of perceptions of the entire relationship, across contracts) and is based on key-informant data. The significant contract-level effects suggest that it is important to examine these decisions at both the firm level (e.g., decision-maker perceptions, relationship characteristics) and the contract level (e.g., service quality, price) to gain a complete understanding of the factors that influence firm upgrade decisions.

Specific insights regarding key factors that influence the customer upgrade decision also emerge from this research. First, we find that a firm's experience of poor service quality can increase the likelihood of upgrading an individual contract. Second, high satisfaction (at the overall firm level) leads to a higher probability of upgrade. Third, customers who pay greater prices for a contract are more likely to upgrade. Finally, we find two key interactions that also influence the upgrade decision: service quality × satisfaction and service quality × price. Next, we discuss each of these findings and elaborate on how service quality influences the upgrade decision through (combined) main and interaction effects.

Poor Service Quality: How Can It Be Good for the Supplier?

At the contract level, poor service quality can increase the likelihood that the firm will upgrade to a higher level of service contract. On the surface, it may seem counterintuitive that by providing poor service, the supplier can gain additional business from the firm. However, it is important to remember that the firm is implicitly making comparisons across contracts; the decision maker evaluates satisfaction with its relationship with the supplier as well as service quality on the focal contract. We speculate that the multidimensional nature of the customer–supplier relationship allows three mechanisms to operate that create a positive main effect of service quality. First, if service quality is poor for the focal contract in an otherwise satisfactory relationship, the decision maker may attribute poor service quality to having purchased a contract level that is insufficient to meet the firm's needs rather than to intrinsically poor service by the supplier. Second, the evaluation of the focal contract quality may lead the decision maker to anticipate regret if he or she repurchases the same level of service contract rather than a higher level of service contract. Third, because customers are committed to buying the service, poor service quality on the focal contract may lead the decision maker to use subtractive framing, whereby he or she realizes that the lower-level service contract is a subset of the higher-level (i.e., augmented) service contract.

Our finding also demonstrates that "snapshots" of contract-level experiences are remembered and that relatively poor service stands out in the mind of the customer in an otherwise satisfactory relationship. It is also consistent with substantial qualitative research showing that critical incidents or moments of truth are important influences on service purchase decisions (Bitner, Booms, and Tetreault 1990) and with prior research on the negativity effect, which suggests that a single negative piece of information carries more weight than a single positive piece of information (Price 1996).

Does this result imply that poor service quality is good for firms? No, certainly not. For this specific firm, the decision to upgrade a contract ultimately depends on the interaction effects. Because most customers are satisfied with the firm and prices are high, the net effect is as follows: The

---

level contracts that are upgraded becomes small in absolute terms. Second, the small sample also leads to a relatively small number of contracts per enterprise. For a random coefficients model, this is problematic because a sufficient number of observations per enterprise (roughly five contracts per customer) is required. Finally, two separate models are less parsimonious than one model with a dummy variable for contract type and (if appropriate) interaction terms.

business customer is less likely to upgrade when quality is poor on the focal contract and the interaction effects are taken into account. We subsequently illustrate this outcome using simulation results.

Higher Prices—More Upgrades?

Firms that currently pay higher prices on a contract (or have lower discounts) are more likely to upgrade. Prior research examining antecedents of the service contract renewal decision (Bolton, Lemon, and Bramlett 2006) has found that price has a negative effect on the renewal decision; the higher the price, the less likely firms are to renew the contract. Thus, it appears that price operates differently on the upgrade decision, perhaps acting as a current reference price against which the increase in price (to move up to a higher-level contract) is considered. This is consistent with recent research on price fairness (Bolton and Alba 2006), which finds that when increases in price are perceived as aligning with increases in vendor costs, such prices are considered fair.

Decision-Maker Perceptions Also Influence the Decision

In addition to effects of contract-level experiences, such as service quality and price, decision makers' perceptions of the supplier, which are firm-level constructs, directly influence firms' contract-level decisions. Examining main effects, we observe that firms with decision makers who are satisfied are more likely to upgrade service contracts. This result is consistent with research in the customer satisfaction and loyalty literature. Notably, we do not find support for a main effect of criticality on the upgrade decision. We believe that this may be a ceiling-effect phenomenon; firms that perceive the equipment as critical to their needs may already have chosen a high level of contract, thereby masking this effect in our data.

Satisfaction Moderates the Effect of Service Quality

Decision-maker perceptions of the supplier moderate the effects of service quality on customer upgrade decisions. Customer satisfaction (the decision maker's evaluation of the supplier's management of the entire account relationship) moderates the effect of service quality on the customer upgrade decision—a rose-tinted-glasses effect. The results suggest that a highly satisfied decision maker can reduce the effects of poor service quality at the contract level, so the firm is less likely to upgrade. A way to interpret this interaction is to consider a firm with low satisfaction (which we believe is driving this result). In this case, it appears that decision makers pay more attention to objective quality information, such as service quality. This implies that they may be in a more active problem-solving mode and is consistent with research by Olsen and Johnson (2003), who find that less satisfied customers focus more immediately on quality perceptions than more satisfied customers. The interplay between supplier perceptions (satisfaction) and contract-level objective experiences of service quality suggests that the firm's decision to upgrade a service contract entails a consideration of information and assessments at multiple levels and that it is more complex than prior research has indicated. Note that from the firm's perspective, our results suggest that decision makers' perceptions can "bias" contract-level service experiences, leading customers to choose contract levels that may not necessarily be optimal.

Price × Service Quality Interaction

In addition to main effects of the contract-level measures of service quality and price, we find that these two factors interact to influence the firm's upgrade decision. This negative interaction effect suggests that, consistent with prior research, firms that experience higher service quality are less price sensitive or, conversely, firms that experience lower service quality are more price sensitive (see Bolton and Myers 2003). This result also implies that there may be a ceiling effect for the positive effects of two somewhat negative aspects of a product (poor quality and high price), such that when the price is too high or the quality is too poor, these positive effects are lessened.

Illustrative Simulation Results

It is important for suppliers to assess the "net effect" of changes in service quality (i.e., main and interaction effects combined) on the buying firm's likelihood of upgrading a service contract. For this supplier, we simulated the effects of a 5% change in satisfaction, resolution time, and price. For each of these three scenarios, we calculated the relative effect on the business customer's probability of upgrading the focal service contract (i.e., not the absolute change in probability). First, a 5% increase in satisfaction leads to a 44% increase in the probability that the business customer will upgrade a contract (ceteris paribus). Second, for a 5% increase in resolution time (i.e., service quality is worse for the focal contract), we find a 25% decrease in upgrade probability. However, if the firm decreases the resolution time by 5% (i.e., service quality is better for the focal contract), the upgrade probability increases by 123%. (This asymmetry arises because there are diminishing marginal returns to service quality.) Finally, for comparison purposes, we simulated the effects of a 5% increase in price. The result is a 6% increase in the upgrade probability (because customers that pay higher prices are less price sensitive as a result of reference price effects). Note that modest service improvements (i.e., average resolution time for the focal contract decreases by approximately two minutes) have a much greater impact on business customer behavior than price changes.

Methodological Contributions

This study's approach to firm decision making is novel in several respects. First, it incorporates perceptual measures (from surveys) and operations measures (from internal process metrics) into a single model of customer behavior. Second, the model is estimated with a cross-sectional, longitudinal database that enables us to show the causal nature of our findings (i.e., antecedents in time $t - 1$ affect the upgrade decision in time t). Third, the inclusion of random effects in the model enables us to account for individual firm effects (i.e., customer heterogeneity) and to distinguish them from the hypothesized effects. Thus, our study shows that comprehensive customer relationship databases can help researchers construct richer models of customer behavior.

Implications for Managerial Practice

A recent study suggests that in most industries, more customers change their spending behavior than defect, implying that gains due to increases in the breadth of the relationship may account for up to 25% of increases in revenue, whereas losses due to defections may account for only 3% of losses in revenue (Coyles and Gokey 2005). This observation is consistent with our findings. Small increases in the likelihood of upgrading are potentially highly lucrative. They yield substantial changes in revenues because there are many customers purchasing relatively high-priced service contracts. Our study offers several insights that enable marketers to manage their relationships with customers. Given the enhanced focus on customer relationships and organic growth, this research is relevant to most industries that provide goods and services in B2B environments, including service firms and traditional manufacturers that now differentiate their offerings by augmenting their tangible goods with service (e.g., delivery guarantees, contracts, consulting). In the following subsections, we offer some guidelines for managers.

Manage at Both the Account and the Contract Level

Suppliers must consider (and manage) business customer experiences simultaneously at the account level (i.e., perceptions that encompass the entire relationship) and at the contract level (e.g., service quality, price, discount) over time. This observation is consistent with prior qualitative research that conceptualizes the evolution of B2B relationships in terms of social exchange processes that arise from discrete episodes or transactions (Narayandas and Rangan 2004).

Understand the Role of Goods and Services

At the account level, suppliers should strive for high customer satisfaction and should also monitor information regarding the objective aspects of the service. The interaction between satisfaction and service quality suggests that account management teams—both sales and service professionals—within the supplier's organization must gain a deep understanding of each business customer's needs so that they can justify to the decision maker the need to upgrade a service contract when satisfaction is low and the service quality is poor (Piercy and Lane 2006). As one decision maker remarked, "[The supplier] needs to be more knowledgeable about our configuration so as to help us more with purchases." In other words, customer dissatisfaction can be an opportunity for account management teams to "migrate" contracts to higher service levels because it may represent a situation in which the decision maker is paying particular attention to the service contracts and is in problem-solving mode. The persuasion process requires an intimate understanding of the environment within the customer organization and how the supplier's product can serve it.

Measure and Manage Service Delivery Metrics

Suppliers must also manage the relationship at the individual contract level. In a B2B context, customers hold many contracts and the contracts are fairly complex, so contract management is a significant challenge. Consequently, suppliers should measure and manage service operations metrics and price levels at the contract level so that account managers are aware of the actual (not promised) service quality levels for every contract the customer purchases and the current price paid on each contract. This knowledge will ensure that appropriate levels of service are being delivered across all contracts. More important, this knowledge will enable the supplier to identify proactively contracts with poor service quality delivery because these contracts represent both a risk and an opportunity. Situations in which the decision maker perceives the supplier's management of the account relationship as satisfactory but objective service quality varies noticeably across contracts simultaneously represent an opportunity for upgrading to a higher-level service contract and a risk of losing the contract. Account management teams should view this situation not only as a critical moment but also as a "teachable moment," when they can discuss an upgrade to a higher-level contract and demonstrate its value (encourage subtractive framing). Similarly, firms that hold contracts that were negotiated at a higher price may be key targets for upgrades because they constitute a market segment that may be more willing to (and interested in) an upgrade.

Manage a Multiple Service Contract Environment

Popular services management heuristics, such as "underpromise, overdeliver" or "a rising tide raises all boats" (i.e., always attempt to exceed customers' quality expectations), are appropriate slogans for encouraging customer retention but not for encouraging upgrading of service contracts. Indeed, these approaches are overly simplistic and difficult to execute in an environment with multiple service contracts. Extremely high service quality on all contracts eliminates the critical moments that suppliers can use to demonstrate the value of a service upgrade for a particular contract. Moreover, if the supplier delivers extremely high service quality for some (but not all) service contracts, the buying firm's expectation or norm is likely to be higher, so contracts with good, but not outstanding, service quality will seem poor in comparison. There may be a tipping point at which rather than upgrading, the buying firm infers that service is unsatisfactory and chooses not to renew the contract.

Improve Customer Decision Making

We must sound a note of caution for managers, such as chief information officers and telecommunication decision makers, who are responsible for managing service contracts. From the buying firm's perspective, our results indicate that a decision maker's perception may color the contract-level experiences, leading him or her to choose contract levels that may not be optimal. It behooves the firm to ensure that the decision maker does not rely solely on his or her impression of how the account is managed. The decision maker must be familiar with individual service contracts, collecting information from the end users of the service so that he or she knows which contracts should be upgraded and which should not.

Overall

In summary, our research suggests that the supplier firm needs to understand the nature and history of customers' experiences with the firm over time, at both the account and the contract levels. Both superior account management and service delivery (through people and technology) are necessary if a supplier firm intends to grow its relationships with customers by means of service contract upgrades. Achieving these synergies may be particularly difficult in firms that isolate marketing, sales, and operations in separate functional silos. Our research is also related to recent research in the area of key account management (KAM). For example, Homburg, Workman, and Jensen (2002) find that activity proactiveness and activity intensity on the part of the supplier firm increase KAM effectiveness. Understanding and monitoring service quality of each contract is a specific example of such proactive activity. In addition, Piercy and Lane (2006) suggest that understanding customer requirements is a key element of KAM effectiveness. The findings of this study suggest specific aspects of customer requirements on which supplier firms should focus. The findings also suggest that suppliers should be asking the following questions about their relationships with their business customers: What is happening in the buyer–seller relationship (i.e., at the account level)? What is happening (operationally) on a contract-by-contract basis? and What is happening when we consider the decision maker's frame of reference (prior satisfaction, perceptions of the environment) and place the focal contract in that context? Answers to these three questions will help suppliers proactively manage their relationships with business customers and enable them to determine which customers may be interested in upgrading and which service contracts they may be willing to upgrade.

Concluding Remarks

There are many avenues for further research regarding upgrading in B2B contexts. First, we conducted our study in only one industry context. However, we believe that this limitation is reduced by the richness of our data, which allows us to eliminate competing explanations for our findings. In a cross-sectional, longitudinal database with multiple contracts at each customer account, we can gain many insights into the factors that influence the upgrade decision. Nevertheless, further research should investigate this decision in different study contexts, as well as experimental settings in which the constructs (e.g., satisfaction, service quality, price) can be manipulated rather than measured. In addition, we incorporated heterogeneity across firms through a random coefficients model. Alternatively, a latent class model could be used to capture unobserved heterogeneity.

Second, we were able to link service operations measures to the customer upgrade decision. However, managers are interested in understanding which improvements in service will "pay off." Thus, further research should link service operations to other business performance metrics, such as customer lifetime value or customer equity. Moreover, it should investigate the effect of service operations improvements on new customer acquisition, retention, upgrading, and cross-buying, perhaps by specifying and estimating systems of equations. For example, understanding the relative influence of service quality variables versus other marketing-mix variables (e.g., price) on these three distinct customer behaviors would also be helpful.

Third, our analyses focus on short-term (i.e., year-to-year decisions) rather than long-term decisions, whereby a customer might choose to stop purchasing any contracts from a given supplier (i.e., completely terminate the relationship). Our model provides insight into the customer upgrade decision over time, but it cannot explain discontinuities, such as when a firm creates a committee to formally evaluate its relationships with suppliers). A worthwhile extension of our work would be to examine switching and upgrade behavior and incorporate transaction costs. It would be helpful to understand the role of competition and brand equity in such decisions, an issue that was beyond the scope of this study.

REFERENCES

Alderson, Wroe (1965), *Dynamic Marketing Behavior: A Functionalist Theory of Marketing*. Homewood, IL: Richard D. Irwin.

Ben-Akiva, M. and S.R. Lerman (1985), *Discrete Choice Analysis*. Cambridge, MA: MIT Press.

Bitner, Mary Jo, Bernard H. Booms, and Mary Stanfield Tetreault (1990), "The Service Encounter: Diagnosing Favorable and Unfavorable Incidents," *Journal of Marketing*, 54 (January), 71–84.

Bloemer, José M.M. and Hans D.P. Kasper (1995), "The Complex Relationship Between Consumer Satisfaction and Brand Loyalty," *Journal of Economic Psychology*, 16 (2), 311–29.

Bolton, Lisa and Joseph Alba (2006), "Price Fairness: Good and Service Differences and the Role of Vendor Costs," *Journal of Consumer Research*, 33 (2), 258–65.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17 (1), 45–65.

——— and James H. Drew (1991), "A Multistage Model of Customers' Assessments of Service Quality and Value," *Journal of Consumer Research*, 17 (4), 375–84.

———, Katherine N. Lemon, and Matthew D. Bramlett (2006), "The Effect of Service Experiences Over Time on a Supplier's Retention of Business Customers," *Management Science*, 52 (12), 1811–23.

———, ———, and Peter C. Verhoef (2004), "The Theoretical Underpinnings of Customer Asset Management: A Framework and Propositions for Future Research," *Journal of the Academy of Marketing Science*, 32 (3), 271–92.

——— and Matthew B. Myers (2003), "Price-Based Global Market Segmentation for Services," *Journal of Marketing*, 67 (July), 108–128.

Boulding, William, Ajay Kalra, and Richard Staelin (1999), "The Quality Double Whammy," *Marketing Science*, 18 (4), 463–84.

———, ———, ———, and Valarie Zeithaml (1993), "A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions," *Journal of Marketing Research*, 30 (February), 7–27.

Brady, M.K., G.A. Knight, J.J. Cronin, G.T.M. Hult, and B.D. Keillor (2005), "Removing the Contextual Lens: A Multi-National, Multi-Setting Comparison of Service Evaluation Models," *Journal of Retailing*, 81 (3), 215–30.

Bridge, R. Gary (2004), "New and Better Managed Services—Maybe," paper presented at Competing Through Service Symposium, Arizona State University (November).

Cannon, Joseph P. and William D. Perreault Jr. (1999), "Buyer–Seller Relationships in Business Markets," *Journal of Marketing Research*, 36 (November), 439–60.

Coleman, J.S. (1990), *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Coughlan, Anne T. (1985), "Competition and Cooperation in Marketing Channel Choice: Theory and Application," *Marketing Science*, 4 (2), 110–39.

Coyles, Stephanie and Timothy C. Gokey (2005), "Customer Retention Is Not Enough," *Journal of Consumer Marketing*, 22 (2), 101–105.

Dickson, Peter (1992), "Toward a General Theory of Competitive Rationality," *Journal of Marketing*, 56 (January), 69–83.

Dodgson, Mark (1993), "Organizational Learning: A Review of Some Literatures," *Organization Studies*, 14 (3), 375–94

Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh (1987), "Developing Buyer–Seller Relationships," *Journal of Marketing*, 51 (April), 11–27.

Folkes, Valerie S. (1988), "Recent Attribution Research in Consumer Behavior: A Review and New Directions," *Journal of Consumer Research*, 14 (March), 548–65.

Ganesh, Jaishankar, Mark J. Arnold, and Kristy E. Reynolds (2000), "Understanding the Customer Base of Service Providers: An Examination of the Differences Between Switchers and Stayers," *Journal of Marketing*, 64 (July), 65–87.

Greene, William H. (2003), *LIMDEP VERSION 8.0: Econometric Modeling Guide*, Vol. 1. Plainview, NY: Econometric Software.

Guadagni, Peter and John D.C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (3), 203–238.

Gustafsson, Anders, Micheal D. Johson, and Inger Roos (2005), "The Effects of Customer Satisfaction, Relationship Commitment Dimensions, and Triggers on Customer Retention," *Journal of Marketing*, 69 (October), 210–18.

Heide, Jan B. and Allen M. Weiss (1995), "Vendor Consideration and Switching Behavior for Buyers in High-Technology Markets," *Journal of Marketing*, 59 (July), 30–43.

Hendrick, Clyde and Susan S. Hendrick (1988), "Lovers Wear Rose Colored Glasses," *Journal of Social and Personal Relationships*, 5 (2), 161–83.

Hoch, Stephen J. and Youn-Won Ha (1986), "Consumer Learning: Advertising and the Ambiguity of Product Experience," *Journal of Consumer Research*, 13 (September), 221–33.

Holbrook, Morris B. (1999), "Introduction to Consumer Value," in *Consumer Value: A Framework for Analysis and Research*, Morris B. Holbrook, ed. New York: Routledge, 1–28.

Homburg, Christian, John Workman Jr., and Ove Jensen (2002), "A Configurational Perspective on Key Account Management," *Journal of Marketing*, 66 (April), 38–60.

Inman, J. Jeffrey and Marcel Zeelenberg (2002), "Regret in Repeat Purchase Versus Switching Decisions: The Attenuating Role of Decision Justifiability," *Journal of Consumer Research*, 29 (June), 116–28.

Jackson, Barbara Bund (1985), "Build Customer Relationships That Last," *Harvard Business Review*, 63 (6), 120–28.

Janiszewski, Chris and Stijn Van Osselaer (2000), "A Connectionist Model of Brand–Quality Associations," *Journal of Marketing Research*, 37 (August), 331–50.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler (1986), "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, 76 (4), 728–41.

Kalwani, Manohar U. and Narakesari Narayandas (1995), "Long-Term Manufacturer–Supplier Relationships: Do They Pay Off for the Supplier Firm?" *Journal of Marketing*, 59 (January), 1–16.

Keaveney, Susan M. (1995), "Customer Switching Behavior in Service Industries: An Exploratory Study," *Journal of Marketing*, 59 (April), 71–82.

Kraljic, Peter (1983), "Purchasing Must Become Supply Management," *Harvard Business Review*, (September–October), 109–117.

Kristof, Amy (1996), "Person–Organization Fit: An Integrative Review of Its Conceptualizations, Measurement and Implications," *Personnel Psychology*, 49 (Spring), 1–49.

Lemon, Katherine N., Tiffany Barnett White, and Russell S. Winer (2002), "Dynamic Customer Relationship Management: Incorporating Future Considerations into the Service Retention Decision," *Journal of Marketing*, 66 (January), 1–14.

Li, Shibo, Baohong Sun, and Ronald T. Wilcox (2005), "Cross-Selling Sequentially Ordered Products: An Application to Consumer Banking Services," *Journal of Marketing Research*, 42 (May), 233–39.

Louviere, Jordan J. (2003), "Complex Statistical Choice Models: Are the Assumptions True, and If Not, What Are the Consequences?" keynote address, Discrete Choice Workshop in Health Economics, University of Oxford (September 22).

Mittal, Vikas and Wagner A. Kamakura (2001), "Satisfaction, Repurchase Intent, and Repurchase Behavior: Investigating the Moderating Effect of Customer Characteristics," *Journal of Marketing Research*, 38 (February), 131–42.

Narayandas, Das and V. Kasturi Rangan (2004), "Building and Sustaining Buyer–Seller Relationships in Mature Industrial Markets," *Journal of Marketing*, 68 (July), 63–77.

Nayakankuppam, Dhananjay and Himanshu Mishra (2005), "The Endowment Effect: Rose-Tinted and Dark-Tinted Glasses," *Journal of Consumer Research*, 32 (December), 390–95.

Ngobo, Paul Valentin (2005), "Drivers of Upward and Downward Migration: An Empirical Investigation Among Theatergoers," *International Journal of Research in Marketing*, 22 (2), 183–201.

Oliver, Richard L. (1997), *Satisfaction: A Behavioral Perspective on the Consumer*. Boston: Irwin/McGraw-Hill.

——— and Raymond R. Burke (1999), "Expectation Processes in Satisfaction Formation: A Field Study," *Journal of Service Research*, 1 (3), 196–214.

Olsen, Line Lervik and Michael D. Johnson (2003), "Service Equity, Satisfaction and Loyalty: From Transaction-Specific to Cumulative Evaluations," *Journal of Service Research*, 5 (3), 184–95.

Park, Whan C., Sung Youl Jun, and Deborah J. MacInnis (2000), "Choosing What I Want Versus Rejecting What I Do Not Want: An Application of Decision Framing to Product Option Choice Decisions," *Journal of Marketing Research*, 37 (May), 187–202.

Piercy, N. and N Lane (2006), "The Underlying Vulnerabilities in Key Account Management Strategies," *European Management Journal*, 24 (2–3), 151–62.

Price, Lydia (1996), "Understanding the Negativity Effect: The Role of Processing Focus," *Marketing Letters*, 7 (1), 53–62.

Rust, Roland T., Jeffrey Inman, Jianmin Jia, and Anthony Zahorik (1999), "What You Don't Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions," *Marketing Science*, 18 (1), 77–92.

———, Anthony J. Zahorik, and Timothy L. Keiningham (1995), "Return on Quality (ROQ): Making Service Quality Financially Accountable," *Journal of Marketing*, 59 (April), 58–70.

Slotegraaf, Rebecca J. and J. Jeffrey Inman (2004), "Longitudinal Shifts in the Drivers of Satisfaction with Product Quality: The Role of Attribute Resolvability," *Journal of Marketing Research*, 41 (August), 260–80.

Stremersch, Stefan, Alan Weiss, Benedict Dellaert, and Ruud Frambach (2003), "Buying Modular Systems in Technology Intensive Markets," *Journal of Marketing Research*, 40 (August), 335–50.

Train, Kenneth E. (2003), *Discrete Choice Methods with Simulation*. Cambridge, MA: Cambridge University Press.

Vancouver J.B. and N.W. Schmitt (1991), "An Exploratory Examination of Person-Organization Fit: Organizational Goal Congruence," *Personnel Psychology*, 44 (2), 333–52.

Verhoef, Peter C. (2003), "Understanding the Effect of Customer Relationship Management Efforts on Customer Retention and Customer Share Development," *Journal of Marketing*, 67 (October), 30–45.

———, Philip Hans Franses, and Janny C. Hoekstra (2001), "The Impact of Satisfaction and Payment Equity on Cross-Buying: A Dynamic Model for a Multi-Service Provider," *Journal of Retailing*, 77 (Fall), 359–78.

Winer, Russell (1986), "A Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research*, 13 (September), 250–56.

Zeithaml, Valarie A. (1988), "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence," *Journal of Marketing*, 52 (July), 2–22.

Pravin Nath & Vijay Mahajan

Chief Marketing Officers: A Study of Their Presence in Firms' Top Management Teams

Not all firms choose to have a chief marketing officer (CMO) in their top management teams (TMTs). This research investigates factors associated with this choice and whether CMO presence/absence in the face of these factors affects firm performance. Findings based on a multi-industry sample of 167 firms over a five-year period (2000–2004) show that innovation, differentiation, branding strategy, diversification, TMT functional experience in marketing, and the chief executive officer being an outsider are associated with the likelihood of CMO presence in the TMT. Furthermore, the authors find that CMO presence in the TMT has neither a positive nor a negative impact on firm performance. The authors discuss the implications of these findings for theory and practice.

Keywords: chief marketing officer, top management team, complexity, power, structure

Over the past three decades, marketing academics have voiced their concern with marketing's decreasing influence at the level of corporate strategy (Anderson1982; Day 1992; Varadarajan 1992; Webster, Malter, and Ganesan 2003; Wind and Robertson 1983). A firm's corporate strategy, which includes major resource allocations, organizational redesign, acquisitions, divestments, and the entering or exiting of major markets, is largely driven by the top management team (TMT) (Cyert and March 1963; Hambrick and Mason 1984; Varadarajan and Clark 1994).[1] Consequently, prior researchers have referred to the presence of the chief marketing officer (CMO) in the TMT as an indicator of both corporate status of marketing and corporate adoption of the marketing concept (Hise 1965; Mann 1971; Piercy 1986; Webster 1981; Webster, Malter, and Ganesan 2003).[2] However, a recent study by Booz Allen Hamilton estimates that only approximately 50% of *Fortune* 1000 firms have a CMO (Hyde, Landry, and Tipping 2004). This is in stark contrast to estimates of more than 80% of firms in *Fortune* 500 having a chief financial officer (CFO) as recently as the year 2000 (Zorn 2004). Little prior research has attempted to explore this phenomenon, a gap in the field that our research aims to fill.

Specifically, the research questions we address herein are, (1) What are the factors associated with the likelihood of CMO presence in firms' TMTs? and (2) What are the consequences of CMO presence for firm performance in the face of these factors? The conceptual model, which we depict in Figure 1 (and explain in the next section), illustrates the thesis for this research with respect to the first research question; that is, the choice of CMO presence/absence is related to the strategic, structural, and environmental factors facing firms' TMTs. Consequently, firm performance implications of CMO presence are contingent on these factors.[3] Such a research endeavor is important given the implications of CMO presence for marketing's role at the level at which corporate strategy is formulated. For example, the CMO has been referred to as the voice of the customer or the customer advocate in the firm's upper echelons (McGovern et al. 2004). In documenting changes in the marketing landscape, Kerin (2005, p. 12) refers to the CMO position as evidence of a "renaissance of sorts" for marketing in "the corporate hierarchy." Crosby and Johnson (2005, p. 13) also suggest that "with the position of CMO, perhaps marketing can bring to the leadership table its strategic role." Therefore, by exploring this phenomenon, our research provides academics and practitioners with an understanding of factors associated with differences in marketing's influence at the level of corporate strategy, adding to the body of knowledge on the role of marketing in the firm. Moreover, it provides a theoretical basis for a mean-

---

[1] The TMT is defined as the executive officers as specified by the firm.

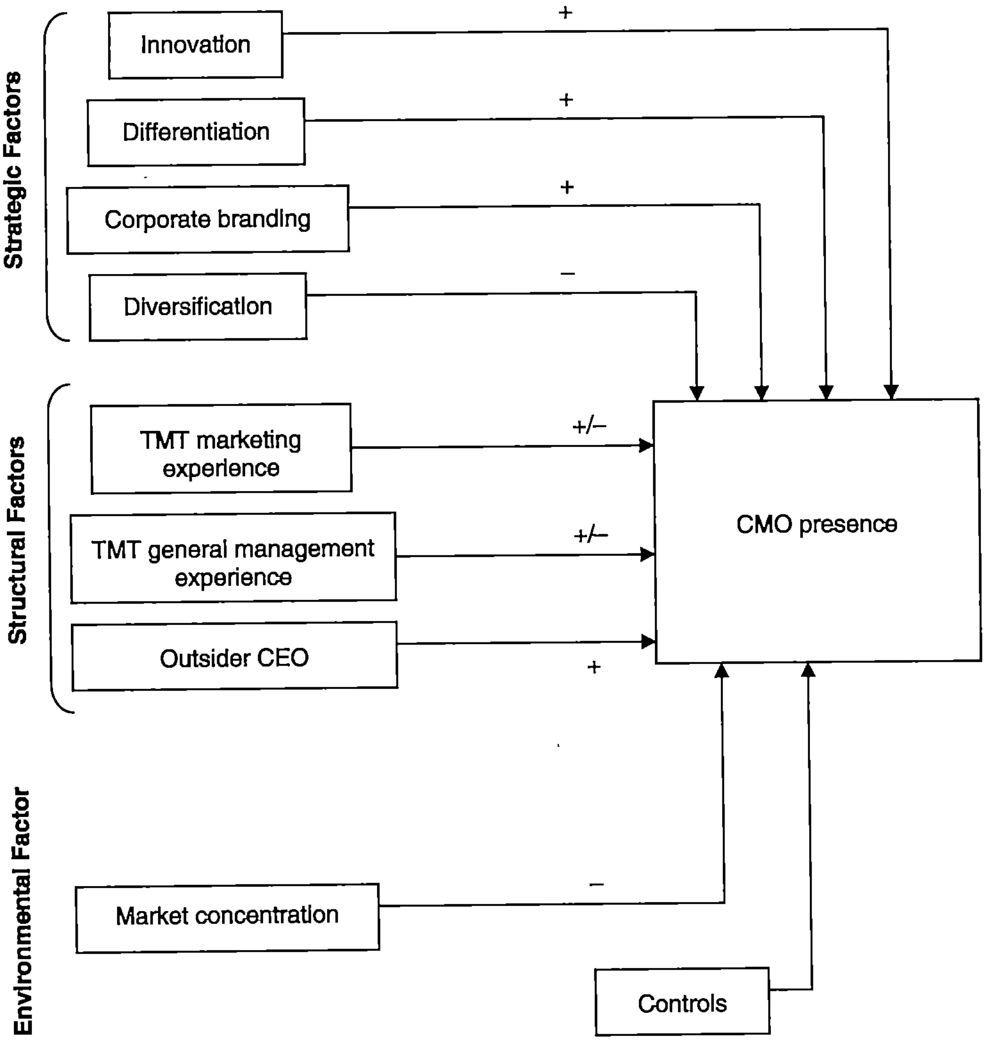
[2] The CMO refers to an executive in the TMT who is explicitly responsible for marketing. Titles of these executives include, but are not limited to, CMO and vice president marketing.

[3] The model for CMO presence implies multiple two-way interactions between CMO presence and the factors associated with it in the model of firm performance. For simplicity, we do not present the latter.

Pravin Nath is Assistant Professor of Marketing, LeBow College of Business, Drexel University (e-mail: pravin.nath@drexel.edu). Vijay Mahajan is John P. Harbin Centennial Chair in Business, Department of Marketing, McCombs School of Business, University of Texas at Austin (e-mail: vijay.mahajan@mccombs.utexas.edu). This research was supported by a Marketing Science Institute Grant and the Dora Bonham Memorial Fund at the University of Texas. The authors thank Raji Srinivasan, Andrew Henderson, Neeraj Bharadwaj, Romana Khan, Rajesh Chandy, Amit Saini, the faculty and doctoral students at the University of Texas at Austin and Drexel University, and the anonymous *JM* reviewers for their insightful comments on this work.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

FIGURE 1
A Model of Factors Associated with CMO Presence in TMTs



Notes: We do not present the conceptual model for the impact of CMO presence on firm performance, given the complexity of two-way interactions suggested by the model of CMO presence and in the interest of space.

ingful dialogue in this regard. Prior research efforts along similar lines have been at the level of the strategic business unit (e.g., Homburg, Workman, and Krohmer 1999; Ruekert, Walker, and Roering 1985). The current research expands the scope of this domain of inquiry to the corporate level, thus providing TMTs with an understanding of firms' central tendencies with respect to the choice of CMO presence. Evidence of practices followed and rationales used by others in industry would help TMTs make better informed decisions with respect to marketing at this level of the firm.

To summarize the findings of our research, firms are more likely to have a CMO in the TMT when they have relatively high levels of innovation and differentiation, when they follow a corporate branding strategy, when the chief executive officer (CEO) is an outsider, and when TMT marketing experience is relatively high. Furthermore, as firms diversify, they are less likely to have a CMO when they are relatively small but are more likely when they are relatively large. The presence of a CMO has no impact on firm performance, and firms with a CMO do not perform better or worse than those without one when faced with the preceding factors. In the following section, we elaborate on the conceptual model for this research illustrated in Figure 1 and present hypotheses along with supporting arguments for both the factors associated with and the consequences of CMO presence. In doing so, we draw on prior literature in marketing and organization theory and on the role of the CMO. Given the scant empirical research on the CMO's role in the firm, this research includes a qualitative stage for this purpose, details of which we also cover in the following section. We test the hypotheses using secondary data. A subsequent section presents the methodology used for this purpose and includes descriptions of the sample and the measures and sources of data for all the variables. This is followed by a section on the analysis of the data and the results. We conclude with a discussion of the results and their implications for marketing theory and practice, along with limitations and future research directions.

Theory and Hypotheses

In this section, we elaborate on the conceptual model for the factors associated with CMO presence illustrated in Figure 1, and we provide arguments for the consequences of CMO presence on firm performance. As the model indicates, we argue that CMO presence is related to a range of factors facing firms. Rationales underlying their choice rely on contingency theory and theories of power dependence and homophily. We begin by addressing each of these theories in the context of the TMT's structural choices. To apply these theories to the choice of a CMO, we then discuss the role of the CMO in the firm. These proceeding subsections then provide the foundation for the choice of factors and hypotheses for each. We conclude this section with a discussion of the hypotheses for firm performance.

Structural Choices in the TMT

Contingency theory suggests that firms seek an optimal fit or match between their structural choices and strategic, structural, and environmental contingencies because some choices may suit some firms better than others (Donaldson 2002; Lawrence and Lorsch 1967; Thompson 1967; Zeithaml, Varadarajan, and Zeithaml 1988). Applications of this theory or framework in the context of TMTs have been to the structural choices made within the TMT in response to conditions that top executives face (e.g., Hambrick and Cannella 2004; Keck and Tushman 1993; Zorn 2004).

Top executives perform multidimensional functions that include monitoring the environment, interpreting information, and making strategic choices (Finkelstein and Hambrick 1996; Hambrick and Mason 1984). At the same time, they are boundedly rational, being constrained by demands on their time and by limited processing abilities (Cyert and March 1963; March and Simon 1958). Consequently, top executives face conditions of complexity and uncertainty. To cope with such conditions, these executives make structural choices or adaptations that aid decision making, especially when these conditions and decisions pertain to critical domains (Hickson et al. 1971). For example, Zorn (2004) finds that uncertainty due to financial reporting requirements explains the increasing prevalence of CFOs in TMTs. As this example indicates, the choice made is either in or closely related to the domain that causes complexity and uncertainty (e.g., a CFO to address complexity in or as a result of the financial domain). The preceding discussion suggests that CMO presence should be related to contingencies that increase TMT complexity and uncertainty in critical aspects of the marketing domain. This inference also relies on the CMO's role being inextricably linked to the marketing domain, for which we provide evidence in the next section.

The Role of the CMO

We identified the CMO's roles on the basis of qualitative research that included analyzing press releases accompanying CMO appointments, CMO job descriptions on company Web sites, and four interviews with senior executives from reputed public, private, and management consultancy firms. Examples of CMO responsibilities included "leading the company's marketing organization"; "responsibility for global marketing resources"; "uniting and strengthening various departments' own marketing plans"; and "directing global marketing efforts, including branding, product marketing, and customer relationship marketing." Not surprisingly, it was evident that the CMO, when present, was focused on the marketing domain. More broadly, it has been found that the function supervised accounts for a large part of the variation in managerial work (Mintzberg 1973). Vorhies and Morgan (2005) find that the following eight capabilities are suitable for benchmarking as marketing capabilities: product development, pricing, channel management, marketing communications, selling, market information management, marketing planning, and marketing implementation. Most descriptions of CMOs' roles included at least some, if not all, of these capabilities.

The scant prior empirical research with respect to the CMO's role also supports the preceding inference of the CMO's role with respect to the marketing domain. The CMO's responsibilities have included advertising, new marketing staff selection, sales forecasting, new product development, marketing research, and marketing planning (Mann 1971; Piercy 1986). Hopkins and Bailey (1984) find that this executive's responsibilities are split between advising corporate management on marketing issues on the one hand and coordinating marketing activities and developing marketing skills at the divisional level on the other hand. More recently, responsibilities of CMOs documented by Hyde, Landry, and Tipping (2004) and McGovern and Quelch (2004) include monitoring the consumer landscape; developing consumer insights; directing brand strategy, marketing planning, and coordination among business units; and managing marketing services such as market research.[4] Finally, normative suggestions on the CMO's role include functions in the marketing domain, such as interpreting extensive market data and providing insights into consumers (Gilliatt and Cuming 1986; Kerin 2005).

CMO-Related Structural Choice in the TMT

The preceding sections lay the foundation for identifying factors related to CMO presence. First, we establish that TMTs address complexity and uncertainty in critical domains by making domain-related structural choices. Second, we provide evidence that the CMO's role is closely related to the marketing domain. Critical elements of the marketing domain include the firm's market-based assets that are made up of relational and intellectual assets with respect to customers and channels (Day 1994; Srivastava, Shervani, and Fahey 1998; Vorhies and Morgan 2005). Therefore, the CMO, when present, is the constituency within the TMT that has significant responsibility for and information about these assets.[5] Armed with this knowledge, the CMO is expected to reduce the complexity and uncertainty that the TMT faces in the marketing domain or in critical decision-making areas affected by the marketing domain. Thus, the factors that are identified to explain the choice of having/not having a CMO are those that affect such marketing-related complexity for the TMT.

---

[4] Prior empirical research has also demonstrated that CMOs can vary in their roles. For example, Piercy (1986) finds that CMOs vary in the extent to which they are responsible for selling, product policy, marketing services, corporate strategy, and physical distribution. Hyde, Landry, and Tipping (2004) distinguish between three emerging CMO models, which, in the order of increasing strategic (versus tactical) responsibility, are marketing service providers, marketing advisors, and drivers of growth. We are unable to distinguish among these models in our research given the use of secondary data. However, we expect that CMOs identified in our study have a primarily strategic role with respect to the marketing domain because unlike these studies, we include as CMOs only those executives listed in the TMT (no such criterion for inclusion is mentioned in these other studies).

At the same time, having a CMO is also tantamount to giving up control over the marketing domain, which other top executives in the firm may be less willing to do. Executive power within the firm draws from control over these resources (Pfeffer 1981). Therefore, the factors that are identified are also those that have a bearing on the power dynamics in the TMT with respect to control over the marketing domain. Finally, because having a CMO has a bearing on the affiliations within the TMT, we also explore the factors linked to homophily, the phenomenon whereby people prefer others who are similar to them (Lazarsfeld and Merton 1954).

Strategic Factors Related to CMO Presence

Innovation and differentiation. Strategies of innovation and differentiation, represented by the emphasis on research and development (R&D) and advertising, respectively, increase the need to identify market opportunities quickly and correctly. Top management is concerned with appropriate resource allocation across a diverse range of new product and positioning initiatives (Finkelstein and Hambrick 1992). This calls for skill and experience in segmentation, targeting, and positioning at the corporate level, or what Hitt and Ireland (1985, 1986) call "corporate-level marketing competencies." Varadarajan and Clark (1994) point out that these are areas in which marketers can contribute to corporate strategy with their experience in business-level segmentation, targeting, and positioning. As firms pursue new product markets and segments, top executives are faced with greater amounts of diverse information. Furthermore, the success of the strategies of innovation and positioning relies heavily on successful development and deployment of marketing capabilities. Therefore, these conditions make it more likely that the TMT will have a CMO to help reduce TMT complexity in the marketing domain while overseeing critical marketing resources.

In support of these arguments, we offer evidence from the qualitative stage of this research; we commonly observed the phrases of "growth," "innovation," and "differentiate" in the job descriptions of CMOs and press releases of their appointments. Formally, we propose the following:

> H_1: A firm's level of innovation is positively related to the likelihood of CMO presence in its TMT.
>
> H_2: A firm's level of differentiation is positively related to the likelihood of CMO presence in its TMT.

Corporate versus other branding strategies. Branding strategies of firms have been classified as corporate, house-of-brands, and mixed (Laforet and Saunders 1994; Rao, Agarwal, and Dahlhoff 2004). A corporate branding strategy uses the corporate brand name for all the firm's products and services, whereas a house-of-brands strategy does not use the corporate brand name at all. A mixed branding strategy has some of the firm's products with the corporate brand name and others with other names. The corporate brand is a critical asset to all firms but more so for firms that rely on it for all their products and services. Therefore, firms with a corporate branding strategy are concerned about monitoring how they build and deploy this asset (Rao, Agarwal, and Dahlhoff 2004). We argue that such a concern makes it more likely for TMTs in firms with a corporate branding strategy to have an executive accountable for the corporate brand. Branding comes under the domain of marketing, and consequently, this executive is the CMO. As one of the top executives interviewed mentioned, "The CMO has to own the metrics area about overall brand health." In firms with either mixed branding or house-of-brands strategies, it is more likely that brands have distinct images. Structurally, there is a greater possibility of senior brand managers who monitor and filter information on their brands to the TMT (Aaker and Jacobson 1994). This layer of senior management makes the task of interpreting complex information on intangible marketing assets easier for the TMT. This argument may suggest that the absence of a CMO in firms without a corporate branding strategy has no bearing on marketing's influence in such firms. However, note that functional representation has been linked to functional importance (Chaganti and Sambharya 1987; Webster, Malter, and Ganesan 2003), and our interest is in the level of the TMT, not the level of middle management, which is where senior brand managers are more likely to be. Therefore, the greater criticality and complexity associated with a corporate branding strategy makes a CMO more likely than not.

Although no prior research has linked branding strategies to criticality and complexity for TMTs, we offer qualitative evidence in support of the arguments. The roles of CMOs were often linked to that of "brand management and marketing," "driving the firm brand," and "brand stewardship (creating a consistent internal and external corporate brand voice)," themes that were recurrent across firms with corporate branding strategies. On the basis of the preceding discussion, we propose the following:

> H_3: The likelihood of CMO presence in the TMT is higher in firms that have a corporate branding strategy than in firms that pursue other branding strategies.

Diversification. As firms diversify into multiple businesses, marketing activities get delegated to the business level (Varadarajan, Jayachandran, and White 2000). Consequently, corporate strategy and its objectives get delineated from business strategy and its objectives. Diversification and the formation of a corporate strategic planning department are structural solutions to firm expansion (Markides 1995). Therefore, we expect that complexity for the TMT decreases as firms diversify. Furthermore, with diversification, heads of business units become independent, powerful constituencies themselves. For them, retaining control over key resources becomes critical to maintaining and increasing power or influence (Doz and Prahalad 1991; Fligstein 1987). Therefore, a CMO is likely to be more valued in single-business firms than in diversified firms. Note that it is not possible to identify which of the preceding two logics is at play, because they both suggest the same direction of relationship. Formally, we propose the following:

---

[5] It could be argued that TMT executives with marketing backgrounds may also play this role. However, regardless of functional backgrounds, they are not explicitly focused on this role, given constraints of their own roles as CEO; COO; or heads of divisions, regions, or other functions (Gilliatt and Cuming 1986).

> H_4: A firm's level of diversification is negatively related to the likelihood of CMO presence in its TMT.

Structural Factors Related to CMO Presence

TMT's functional background in marketing. Prior research has viewed functional experience or expertise in the TMT as a structural response to contingencies facing firms (Chaganti and Sambharya 1987; Keck and Tushman 1993). The type of functional experience in the TMT is viewed as relevant in addressing TMT complexity. Furthermore, it has been shown to influence power dynamics within the TMT (Hambrick 1981). Therefore, we explore the role of the functional experience in the TMT as a factor related to CMO presence. Given that general management experience involves overseeing activities of multiple functions, including marketing, we offer similar hypotheses for both these types of functional experience. For example, Day and Nedungadi (1994) suggest that the managerial representations of marketing managers and general managers are likely to be somewhat similar.

Two lines of reasoning suggest competing or alternative hypotheses. The first is the phenomenon of homophily, according to which people tend to like those who are similar to them (Lazarsfield and Merton 1954; McPherson and Smith-Lovin 1987). Functional experience has been shown to discriminate fairly well between people (Dearborn and Simon 1958; Hambrick and Mason 1984). Homophily would then suggest that greater marketing or general management experience in the TMT would be related to a greater likelihood of a CMO. Related to the idea of homophily is that top executives with marketing or general management experience also recognize the importance of the voice of the customer in the TMT. For example, Webster, Malter, and Ganesan (2003) find this to be the case for firms with marketing CEOs. Therefore, we propose the following hypotheses for marketing and general management experience:

> H_5: The proportion of a firm's TMT with marketing experience is positively related to the likelihood of CMO presence in its TMT.
>
> H_6: The proportion of a firm's TMT with general management experience is positively related to the likelihood of CMO presence in its TMT.

However, a second line of reasoning would suggest an opposite effect. Similar functional backgrounds are associated with similar mental maps (Dearborn and Simon 1958; Hambrick and Mason 1984). Greater experience in the TMT in marketing or general management could lead to an overlap of cognitive strategies. Thus, from a contingency point of view, a CMO may not be required if the TMT has executives with marketing or general management expertise. In addition, a CMO's similar expertise may be viewed as a challenge to the other executives in the TMT (Pfeffer 1981). Note that both rationales, one that draws on the contingency view and the other that draws on power dependence, have the same effect. On the basis of the preceding discussion, we propose the following competing hypotheses for marketing and general management experience:

> H_{5Alt}: The proportion of a firm's TMT with marketing experience is negatively related to the likelihood of CMO presence in its TMT.
>
> H_{6Alt}: The proportion of a firm's TMT with general management experience is negatively related to the likelihood of CMO presence in its TMT.

CEO's firm-specific operational experience (outsider versus insider CEOs).[6] Whereas the preceding discussion viewed the entire TMT's functional experience, it is also important to focus specifically on the CEO, given his or her integrative role within the TMT (Calori, Johnson, and Sarnin 1994). The CEO has multiple responsibilities (Mintzberg 1973), and we expect that in an effort to discharge these effectively, the CEO attempts to address his or her own limitations. A CEO limitation that has been relevant in explaining TMT staffing decisions is the CEO's first-hand experience with operational activities specific to the firm (Cannella and Lubatkin 1993). For example, Hambrick and Cannella (2004) find that CEOs who lack this experience, also referred to as outsider CEOs, are more likely to have chief operating officers (COOs) than insider CEOs, who have risen through the ranks of the firm and are more familiar with the firm's operations. This logic would be relevant to all the functions at the operational level, and we extend it to the marketing domain. Thus, we hypothesize the following:

> H_7: The likelihood of CMO presence in the TMT is higher in firms with an outsider CEO than in firms with an insider CEO.

Environmental Factors Related to CMO Presence

As market concentration increases, industries tend to be dominated by fewer and larger firms. Therefore, the diversity of possible actions under the marketing domain is relatively greater in industries with low levels of concentration. For example, Kuester, Homburg, and Robertson (1999) find that breadth of retaliatory behavior, in terms of the number of marketing instruments used, is greater in less concentrated industries. Such industries also tend to be associated with customer-oriented managerial representations rather than competitor-oriented ones that are found in highly concentrated industries (Day and Nedungadi 1994) because there are many bases of differentiation in the former, leading to a greater need to focus on customer needs, whereas a few large players drive the market in the latter (Porter 1980). Thus, it is more likely that there is greater complexity for the TMT in the marketing domain in less concentrated industries. This would suggest a greater likelihood of CMO presence in such industries. Note that some empirical evidence also suggests that CMOs are less likely in markets with relatively low concentration. Anderson, Fornell, and Mazvancheryl (2004) find that customer satisfaction's positive impact on shareholder value is weaker in less concentrated industries. Their rationale is that given the intense rivalry in such industries, customer satisfaction is not enough to hold back consumers (Kohli and Jaworski 1990). It is possible to stretch this rationale and argue that CMOs are less valued in less concentrated industries, given that a key marketing asset is relatively less critical. However, for this rationale to affect CMO presence, firms must make CMO-related choices in response to the stock market's reaction, for which there is no evidence. Overall, the arguments discussed in this section seem to suggest that low, rather than high, levels of market concentration are associated with the likelihood of CMO presence. Formally, we propose the following:

---

[6] We thank an anonymous reviewer for suggesting a thorough exploration of the CEO's role.

> H_8: The degree of market concentration in a firm's primary industry is negatively related to the likelihood of CMO presence in its TMT.

Consequences of CMO Presence

The preceding sections argue that CMO presence/absence is contingent on the level of the strategic, structural, and environmental factors facing firms. The logic of contingency theory suggests that the impact of CMO presence on firm performance depends on the levels of these factors (Lawrence and Lorsch 1967; Zeithaml, Varadarajan, and Zeithaml 1988). To the extent that firms face conditions that argue for CMO presence/absence, we expect that CMO presence/absence helps such firms improve performance. Conversely, when firms make this choice in opposition to the logic of the conditions they face, we expect that CMO presence/absence hampers firm performance. We hypothesize these interaction effects for all but two of the factors associated with CMO presence. It is not possible to offer similar hypotheses for TMT marketing and general management experience, because alternative logics suggested competing hypotheses for the effects of these factors on CMO presence. Therefore, we propose the following:

> H_9: Firm performance is improved by CMO presence in the TMT for firms pursuing (a) relatively high levels of innovation, (b) relatively high levels of differentiation, (c) a corporate branding strategy, and (d) relatively low levels of diversification, as well as for firms (e) with an outsider CEO and (f) in industries that are relatively less concentrated.

Our conceptual framework for the explanation of CMO presence does not offer any guidance regarding the main effect of the presence or absence of a CMO on firm performance. Recent advocates of CMO presence suggest that having such an executive raises the importance of marketing issues at the corporate level and can only benefit firm performance (Kerin 2005; McGovern et al. 2004). At the same time, the presence of such an executive also adds administrative and bureaucratic costs (Taylor 2004). The only evidence we found in this regard was an inferred indirect association of TMT marketing experience with firm performance through sales growth (Weinzimmer et al. 2003). Therefore, we do not offer a hypothesis for the main effect of CMO presence on performance. However, we explore whether firms that have CMOs perform better or worse than firms that do not.

Methodology

Sample

We chose to observe firms over a considerable duration rather than at any one point in time so that the results did not reflect conditions or reporting idiosyncrasies of a particular year. At the time this research was initiated, the most recent year with secondary data available on TMTs of firms was 2004. Consequently, we chose to observe firms over the five-year period from 2000 to 2004. All firms in the COMPUSTAT database with sales of at least $250 million in 2002, the midpoint of our period of observation, were identified. We believed that limiting the research to larger firms was an empirical necessity given the low prevalence of CMOs even in the *Fortune* 1000, as observed in the Booz Allen Hamilton study (Hyde, Landry, and Tipping 2004). Furthermore, Hambrick and Cannella (2004) use a similar cutoff. From this set, we retained only firms without missing data on the various factors. To test the hypothesis on innovation, we dropped firms in industries that did not report R&D expenditures (e.g., retailing). As a result of these filters, the final sample had 167 firms that represented a cross-section of industries.[7]

The percentage of firms with a CMO was approximately 40% each year—specifically, 41%, 42%, 44%, 41%, and 39% in 2000, 2001, 2002, 2003, and 2004, respectively. These proportions were not significantly different from one another. Approximately 25% of the firms had a CMO, whereas 41.32% did not in all the years of observation. An additional 17.37% of the firms had a CMO, whereas 16.17% did not for more than half of the period of observation. We also note that 19.6% of the CMOs across all the years in the sample had the actual title of CMO, and the trend for the title increased over time. As an aside, 97.2% of the firms' TMTs had a finance executive (of these, 96.8% had the title of CFO) across all the years of observation.

---

[7] The break-up of the sample by industry (two-digit Standard Industrial Classification [SIC]) was 38 firms in Business Services (SIC 73), 37 in Industrial Machinery and Equipment (SIC 35), 29 in Electrical and Electronic Equipment (SIC 36), 21 in Chemicals and Allied Products (SIC 28), 19 in Instruments and Related Products (SIC 38), 7 in Rubber and Miscellaneous Plastic Products (SIC 30), 6 in Fabricated Metal Products (SIC 34), 4 in Furniture and Fixtures (SIC 25), 3 in Paper and Allied Products (SIC 26), and 3 in Primary Metal Industries (SIC 33).

Data Sources and Measures

We collected data using secondary sources. Unless specifically mentioned, we collected most measures annually over the 2000–2004 period. We now describe the measures and sources of data for the various dependent and independent variables.

Operationalization of the TMT. Prior researchers have operationalized the TMT in various ways, depending on the research purpose, methodology used, and data availability (for a list of these operationalizations, see Table 1 in Carpenter, Geletkanycz, and Sanders 2004, pp. 754–58). These operationalizations include vice president or senior vice president and above (Chaganti and Sambharya 1987; Keck and Tushman 1993; Michel and Hambrick 1992; Weinzimmer et al. 2003), the top two levels (Wiersema and Bantel 1992), and inside directors (Finkelstein 1992). A few researchers have also used key informants such as the CEO to identify the TMT (e.g., Bantel and Jackson 1989). We use the most inclusive of operationalizations that allow for consistency across firms and with more recent research (e.g., Hambrick and Cannella 2004). Specifically, in this research, the TMT is the list of executive officers specified by a firm in the 10-K or proxy, its annual filings with the Securities Exchange Commission. The mean size of the TMT in the sample of firms was 9.65, and the standard deviation was 3.71, which is consistent with prior research using broader or more inclusive operationalizations.

Dependent variable: CMO presence. We used 10-Ks or proxy statements to identify the dependent variable of interest—the presence or absence of a CMO in the firm's TMT, coded as 1 or 0, respectively. An executive in the TMT with the term "marketing" in his or her title constitutes CMO presence; a TMT without such an executive represents CMO absence. The actual titles included vice president marketing, senior vice president marketing, or executive vice president marketing. We encountered a few cases in which the title had the terms "branding" or "corporate marketing." The title of CMO was typically in addition to these other titles (McGovern and Quelch 2004). We also checked CMO presence coded from the 10-K/proxy against the list of officers provided in annual reports.

Independent variables: strategic, structural, and environmental factors. All measures and data sources have been used in prior research. Innovation and differentiation were captured by the ratios of R&D to sales and advertising to sales, or R&D intensity and advertising intensity, respectively (Bettis and Mahajan 1985). These data are available through Standard & Poor's COMPUSTAT. A graduate research assistant and one of the researchers separately coded the branding strategies of firms, drawing on Rao, Agarwal, and Dahlhoff's (2004) definitions. If the firm had a corporate branding strategy, this variable was coded as 1; otherwise, the variable was coded as 0. Four sources of information—Competitive Media Reports, Datamonitor Reports, 10-Ks, and company Web sites—were used for this purpose. The degree of agreement between the two coders was 80.5%. After discussion, the degree of agreement improved to 93.4%, with the remaining cases of disagreement being assigned codes by the researcher. In line with Rao, Agarwal, and Dahlhoff, we assumed that the firm's branding strategy was constant over the observation period. We computed diversification using the entropy measure based on four- and two-digit-level segment sales available through COMPUSTAT (Palepu 1985). We calculated measures of total, related, and unrelated diversification and tested the hypotheses on diversification using each separately. A large proportion of the sample had diversification measures of zero by this method. Consequently, we also used a dummy variable coded as 0 or 1 to indicate nondiversified versus diversified firms, respectively, for each measure of diversification.

To determine the proportion of the TMT with marketing and general management experience and whether a CEO was an outsider or an insider, we collected biographical information on the executives using multiple sources. These included the 10-Ks, *Bloomberg*, *Dun & Bradstreet's Register of Corporations*, and company Web sites. The method of classification of functional experience is fairly well established in prior research (for a review, see Finkelstein and Hambrick 1996). Functional experience in marketing and/or marketing-related functions that included sales, advertising, brand management, and customer service was coded as marketing functional experience. Experience as a head of a division or region or as a top management executive without any specific functional responsibilities was coded as general management experience. The coding was fairly straightforward and was carried out by graduate research assistants, who we trained. Some information was publicly available on approximately 95% of the executives in the TMTs of the firms in the sample. The mean and median of years of available biographical information on these executives was 13.88 and 12, respectively. When no information was available on an executive, we dropped the executive from the TMT for the purpose of analysis. Consequently, 75% of the firms had some biographical information on all executives in the TMT, and 95% of the firms had some biographical information on at least 75% of the executives in the TMT. The measures we used were either the proportion of the executives in the TMT with experience in marketing or general management or the proportion of the total years of experience in the TMT in marketing or general management, after we excluded the CMO from the TMT.[8] As an alternative to using the entire TMT for determining marketing experience, we used a dummy measure that equaled 1 when the CEO of the firm had marketing experience. Recent research by Webster, Malter, and Ganesan (2003) suggests that CMOs are more likely and marketing is more valued in firms in which CEOs have a marketing background. A dummy variable for outsider (insider) CEO was coded as 1 (0) if the executive had spent zero (more than zero) years in the firm before being appointed as a CEO. We used other cutoffs in line with prior research on outsider

---

[8] We thank an anonymous reviewer for pointing out the importance of excluding the CMO while calculating the TMT's experience in marketing and general management

CEOs (e.g., Cannella and Lubatkin 1993), but the results did not differ across alternatives. We also used the number of years instead of the preceding categorical operationalization, but we do not report the results, because they were significantly weaker.

We calculated the environmental factor of market concentration using the Herfindahl–Hirschmann index (HHI), which is the sum of the square of market shares at the two-digit Standard Industrial Classification (SIC) level, of all firms listed in COMPUSTAT. As alternatives to the HHI, we also used the four-firm and eight-firm concentration ratios, which add the market shares of the four and eight firms with the largest market shares, respectively, at the two-digit SIC level.

Dependent variable for model of consequences: firm performance. We used various measures of firm performance that have been used extensively in prior research—namely, sales growth, a market-based measure of a firm's performance that, with adequate controls, captures the notion of market share; return on assets (ROA) and return on sales (ROS), which are accounting-based performance measures; and Tobin's q, a market-based measure that reflects the value of a firm as perceived by its shareholders. We calculated sales growth as the increase in sales as a proportion of the sales in the preceding year. We calculated ROA as the ratio of profits to assets and ROS as the ratio of profits to sales. We calculated Tobin's q using Chung and Pruitt's (1995) approximation—namely, the ratio of the sum of market value of the firm and book value of its debt to its total assets. Researchers prefer Tobin's q because it is forward looking and is not affected by accounting standards that may differ across industries (for details, see Anderson, Fornell, and Mazvancheryl 2004, p. 183). We obtained all required data for the calculation of these variables from COMPUSTAT.

Controls

Because the sample of firms is drawn from multiple industries, we controlled for industry effects using median measures for the variables at the two-digit SIC level to the extent that data were available (Hambrick and Cannella 2004).[9] These were subtracted from the firm's measure on that variable. For R&D intensity, advertising intensity, sales growth, ROA, ROS, and Tobin's q, we used all firms in each SIC with available data in Standard & Poor's COMPUSTAT for this purpose. As an alternative to industry medians, we also used dummy variables for industries at the two-digit SIC level. In addition, we controlled for the size of the firm measured through the natural log of the number of employees reported in COMPUSTAT. Firm size controls for explanations that draw on the tendency of leading firms to be initiators of structural choices, such as having a CMO; following prior research, we also control for size in the model of performance. We also controlled for CEO tenure; the presence/absence of a COO, which we coded as a dummy variable (1/0); and the extent of divisionalization in the firm, which we measured as the proportion of executives in the TMT who were divisional heads. We obtained these measures from the 10-Ks. We included CEO tenure to control for the possibility of CMO presence/absence being part of a TMT shake-up that usually accompanies new CEO appointments. We needed to control for COO presence because it has been shown to be related to outsider CEOs for reasons similar to those that we hypothesized for CMO presence (Hambrick and Cannella 2004). We controlled for divisionalization because a preponderance of such executives may imply CMOs at the divisional level rather than at the corporate level. We also included prior performance, which we measured as sales growth, ROA, or ROS, as a control measure in the event that CMO presence/absence was a response to either poor or superior performance. Finally, we controlled for the consumer profile of the firm, in terms of whether the firm was a pure business-to-consumer player, a pure business-to-business one, or one that sold to both ultimate and organizational consumers (mixed) because the propensity to have a CMO may differ across these contexts. We coded firms as one of these according to the description of the firm's business and its customers as reported in the 10-K. In line with Rao, Agarwal, and Dahlhoff (2004), we assumed that the firm's consumer profile was constant over the observation period. Only 9% of the firms were pure business-to-consumer firms. Therefore, we combined business-to-consumer and mixed firms and coded them as 1; we coded business-to-business firms as 0. In the model of firm performance, in addition to most of the preceding control variables, we controlled for relevant performance variables by drawing on prior research; for example, with Tobin's q as a measure of performance, we controlled for prior profitability and sales growth because these have been shown to affect this measure of performance.

Analysis and Results

Table 1 presents descriptive statistics and correlations for all measures, pooled over the period of observation. None of the correlations exceeded .5. For all models discussed, none of the variation inflation factors were above 4, and none of the condition indexes associated with the eigenvalues of the variable matrix exceeded 15. These tests imply the absence of significant multicollinearity problems (Johnston 1991, p. 250). Longitudinal data enabled us to perform two types of analyses for testing the hypotheses explaining CMO presence. In the first analysis, we assumed that the decision to have a CMO is revisited every year. Here, we modeled CMO presence in 2001–2004 as a function of the hypothesized variables and controls; those that varied over time were lagged by a year, except for firm size, outsider CEO, CEO tenure, COO presence, and divisionalization, which were not lagged. Empirically, lagging the time-varying independent variables enabled us to rule out reverse causality, especially in the case of the strategic factors. We also

---

[9] Because the distributions were skewed, it is more appropriate to use the median, not the mean, as a measure of central tendency or as a control (Hambrick and Cannella 2004; Rao, Agarwal, and Dahlhoff 2004).

TABLE 1
Descriptive Statistics and Correlation Coefficients

	M (SD)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. CMO presence$_t$	.42 (.50)															
2. Innovation$_{t-1}^{a}$	−.03 (.09)	.20***														
3. Differentiation$_{t-1}^{a}$	.01 (.03)	.16***	−.05													
4. Corporate branding^b	.53 (.50)	.26***	.31***	−.09**												
5. Total diversification$_{t-1}^{c}$	.30 (.42)	−.09	−.15***	.06*	−.33***											
6. TMT marketing experience$_{t-1}^{d}$	.15 (.14)	.14***	.07*	.15***	.03	−.15***										
7. TMT general management experience$_{t-1}^{d}$	.53 (.18)	−.14***	−.13***	.03	−.10***	.04	.04									
8. Outsider CEO$_t$	.32 (.47)	.14***	.01	.12***	.05	.04	−.05	−.01								
9. Market concentration$_{t-1}^{e}$	.04 (.02)	−.00	.20***	.03	−.06	.16***	.01	−.01	.15***							
10. Log(number of employees)$_t$	8.65 (1.37)	−.06	−.15***	.07*	−.32***	.43***	−.07*	.17***	−.05	.08**						
11. Year	2002.5 (1.12)	−.02	.08*	.04	.00	.03	.04	−.05	−.00	.11**	.01					
12. CEO tenure$_t$	6.72 (7.71)	−.08**	−.07*	−.12***	.11***	−.13***	−.04	.05	−.05	−.08**	−.10**	.02				
13. COO presence$_t$	.29 (.45)	−.08**	−.09**	.08*	.06	−.07*	.03	.05	.03	−.12**	.05	−.03	.13***			
14. TMT divisionalization$_t$	.19 (.17)	−.06*	.04	.02	−.08**	.21***	.08**	.27***	.02	.06	.19***	.03	−.08**	−.09**		
15. Prior performance$_{t-1}^{f}$	.06 (.35)	.04	.04	.12**	.07*	−.13***	.10**	−.12	.00	−.05	−.08**	−.13***	−.02	.07*	−.08**	
16. Business-to-consumer/mixed versus business-to-business firm^b	.43 (.49)	−.07*	−.24***	.43***	−.42***	.19***	.13**	.04	.08**	.06	.26***	.00	−.18***	.02	.05	−.03

*$p < .10$.
**$p < .05$.
***$p <$ 01.
[a] Ratios of R&D (innovation) and advertising (differentiation) to firm sales, respectively, less industry median ratios.
[b] Assumed to be constant over the period of observation.
[c] Entropy measure of total diversification.
[d] Proportion of executives in the TMT (excluding the CMO from time t if present in the TMT in time t − 1) with marketing or general management experience.
[e] HHI
[f] Sales growth less industry median values.
Notes: Table shows correlations between measures pooled across 167 firms and four years. N = 668 firm years.

included the year as an additional control in case CMO presence varied over time. Given that our dependent variable of CMO presence is binary and that we repeatedly observe the same firms over a period of time, we used the generalized estimating equations (GEE) approach (Zeger and Liang 1986). The GEE approach estimates regression coefficients and standard errors after weighting data with estimates of serial correlations that occur with repeated observations. This approach also allows for the specification of various distributions for the dependent variable and, consequently, has been used to model binary outcomes in prior research (Hambrick and Cannella 2004; Law 2002). The results from this analysis, using StataCorp (2005), are shown as Models 1 and 2 in the first two columns of Table 2; Model 2 introduces two-digit SIC dummy variables instead of using industry-level variables as controls.

In the second analysis, we study a firm's propensity for a CMO over the entire period of observation. Here, for a CMO to be considered present/absent in a firm, the firm should have had/not had this executive in the TMT for a significant portion of the period of observation. In the interest of space, we report the results only for CMO presence/absence, when firms had/did not have a CMO in the TMT for more than half the period of observation. The results are fairly similar across other specifications, except that the usable sample for analysis decreases in size; for example, the sample reduces from 167 to 106 when we include only those firms that have/do not have a CMO throughout the period of observation. We averaged time-varying independent variables over the period of observation, and the control of prior performance refers to the performance in the year before the period of observation. We provided this second analysis as a support for the results from the longitudinal analysis, especially in the event that the decision to have a CMO is not an annual one. The results from this analysis, which employs logistic regression, are shown as Model 3 in the last column of Table 2. The analysis that tests the hypotheses related to consequences of CMO presence on firm performance also uses averages to avoid the complexity of trying to interpret short-term effects of CMO presence and separate cause and effect. The results from this analysis, which uses ordinary least squares, appear in Table 3. In the interest of space, we present the results using only two of the measures of performance—Tobin's q in Model 1 and sales growth in Model 2. The results were not significantly different when we used ROA and ROS. In general, the preceding analyses follow the work of Hambrick and Cannella (2004), who study factors associated with COO presence and the consequences of COO presence for firm performance.

Results of the Model of Factors Associated with CMO Presence

As Table 2 shows, we find support for the hypothesized positive association between innovation and CMO presence (H_1) ($p < .05$ in Models 1 and 3; $p < .10$ in Model 2). H_2, which predicted the positive association between differentiation and CMO presence, is also strongly supported ($p < .01$ in Model 2; $p < .05$ in Models 1 and 3). As H_3 predicted, CMO presence is more likely in firms with a corporate branding strategy ($p < .01$ in Model 1; $p < .05$ in Model 2; $p < .10$ in Model 3). We do not obtain a main effect for diversification, in line with H_4. However, post hoc, we find a significant interaction with firm size, after mean-centering both variables ($p < .05$ in Model 2; $p < .10$ in Models 1 and 3). This interaction suggests support for H_4; that is, in relatively small firms, there is a decrease in likelihood of CMO presence as diversification increases. In relatively large firms, we observe an effect in the direction opposite to what we hypothesized. We illustrate this interaction effect in Figure 2 and discuss it in detail in the next section. We find partial support for the hypothesized effects of TMT marketing and general management experience. These variables are significant only in Model 3. We presented competing hypotheses for these structural factors. In line with H_5, CMO presence is more likely as TMT marketing experience increases ($p < .05$). Conversely, CMO presence is more likely as the general management experience in the TMT decreases, in weak support of H_{6Alt} ($p < .10$). We find strong support for H_7, which predicted that outsider CEOs would be more likely to have CMOs ($p < .01$ in Model 3; $p < .05$ in Models 1 and 2). We did not find support for H_8, which offered predictions about the environmental factor of market concentration. Note that we obtain similar results with all the alternative measures of market concentration. The partial support for TMT experience is not achieved when we use the alternative measure of the years of experience in marketing and general management. The results for total diversification are not replicated with measures of related and unrelated diversification. In the model that uses dummy variables to capture industry effects, none of the dummy variables were significant. Finally, a CEO with a marketing background was not significantly related to CMO presence.

Results of the Model of Consequences of CMO Presence

As Models 1a and 2a in Table 3 show, we find no main effect for CMO presence/absence on firm performance.[10] We introduced separately all interactions between CMO presence and the various factors associated with CMO presence. None of these interactions were significant. In the interest of space, we do not present all these nonsignificant interactions, but we demonstrate one of them in Models 1b and 2b of Table 3. Thus, H_{9a}–H_{9f} were supported.

Discussion and Implications

The purpose of this research was to address the following two questions: (1) What are the factors associated with the likelihood of CMO presence in firms' TMTs? and (2) What are the consequences of CMO presence for firm performance in the face of these factors? We were able to identify a mix of strategic and structural factors that are associated

---

[10] We checked for any possible endogeneity by introducing the residual from a logistic regression of CMO presence on variables from the model of CMO presence and from the model of performance into the model of performance. The estimate for this residual was not significant, suggesting that endogeneity was not an issue (Woolridge 2003).

TABLE 2
Results of Logistic Regressions with CMO Presence as the Dependent Variable

Independent Variables	Model 1[a]		Model 2[b]		Model 3[c]		Support for Hypotheses Across Models 1, 2, and 3
Constant	138.95	(125.56)	55.45	(119.98)	–.14	(.74)	—
Innovation[d] (H_1)	2.79	(1.51)**	2.52	(1.49)*	5.97	(2.56)**	Yes
Differentiation[d] (H_2)	12.62	(5.12)**	13.18	(5.07)***	15.34	(7.75)**	Yes
Corporate branding (H_3)	.93	(.33)***	.87	(.37)**	.75	(.45)*	Yes
Diversification (H_4)	–.45	(.42)	–.39	(.44)	–.47	(.61)	No (main effect)
TMT marketing experience (H_5/H_{5Alt})	–.25	(.71)	–.42	(.70)	3.33	(1.55)**	Partial (H_5)
TMT general management experience (H_6/H_{6Alt})	–.34	(.58)	–.43	(.59)	–2.06	(1.24)*	Partial (H_{6Alt})
Outsider CEO (H_7)	.49	(.19)**	.47	(.19)**	1.27	(.46)***	Yes
Market concentration (H_8)	–8.72	(6.70)	—		–9.90	(9.18)	No
Firm size	.03	(.12)	.07	(.12)	.24	(.17)	—
Diversification × firm size (H_4: post hoc)	.43	(.21)*	.44	(.22)**	.58	(.33)*	Yes (interaction effect)
Year	–.07	(.06)	–.03	(.06)	—		—
CEO tenure	–.01	(.01)	–.01	(.01)	–.02	(.03)	—
COO presence	–.12	(.21)	–.10	(.21)	–.49	(.55)	—
TMT divisionalization	.25	(.53)	.20	(.53)	–1.50	(1.44)	—
Prior performance[d]	–.05	(.21)	–.05	(.21)	–.15	(.24)	—
Business-to-consumer/mixed versus business-to-business firm	–.06	(.36)	–.02	(.36)	–.52	(.49)	—
Nine industry dummies[e]	—		Included		—		—

*$p < .10$.
**$p < .05$.
***$p < 01$.

[a] Pooled logistic regression using GEE over the 2001–2004 period. We controlled for industry effects for relevant strategic variables by subtracting the median value at the two-digit SIC level from the raw value of the variable. We coded CMO presence as 1 (0) for firms with (without) a CMO in time t; we measured all time-varying independent variables in time t – 1, except for outsider CEO, firm size, COO, CEO tenure, and divisionalization, which we measured in time t (N = 645 firm years; CMO presence = 42.05%; $\chi^2(16) = 45.73$, $p < .001$).

[b] Model 2 is similar to Model 1, except that we controlled for industry effects using dummy variables at the two-digit SIC level (N = 645 firm years; CMO presence = 42.05%; $\chi^2(24) = 50.22$, $p < .01$).

[c] Logistic regression with CMO presence was coded as 1 (0) for firms with (without) a CMO for more than half the period of observation of 2000–2004, with time-varying independent variables averaged over this period (N = 167; CMO presence: 42.52%; Model $\chi^2(15) = 44.51$, $p < .001$; Nagelkerke $R^2 = .31$).

[d] In Models 1 and 3, innovation, differentiation, and firm performance are the raw values less the median values at the two-digit SIC level.

[e] None of the industry dummy estimates were significant.

Notes: Table shows parameter estimates along with standard estimates in parentheses.

TABLE 3
Results of Ordinary Least Squares Regressions with Firm Performance as the Dependent Variable

Independent Variables	Dependent Variable: Tobin's q[a]		Dependent Variable: Sales Growth[a]	
	Model 1a[b]	Model 1b[b]	Model 2a[c]	Model 2b[c]
Constant	.59 (.13)***	.60 (.13)***	.04 (.02)**	.04 (.02)**
Innovation[a]	5.56 (1.00)***	5.52 (1.00)***	.11 (.13)	.12 (.13)
Differentiation[a]	10.25 (2.82)***	10.59 (2.86)***	.56 (.37)	.52 (.37)
Corporate branding	–.08 (.18)	–.07 (.17)	–.01 (.02)	–.01 (.02)
Diversification	–.20 (.23)	–.20 (.23)	–.08 (.03)***	–.08 (.03)***
TMT marketing experience	–.48 (.65)	–.53 (.66)	.02 (.08)	.03 (.08)
TMT general management experience	.33 (.46)	.35 (.46)	.02 (.06)	.01 (.06)
Outsider CEO	–.59 (.18)***	–.46 (.25)*	.01 (.02)	–.01 (.03)
Market concentration	–.16 (3.27)	–.13 (3.27)	–.66 (.45)	–.66 (.45)
Firm size	.23 (.07)***	.23 (.07)***	–.00 (.01)	–.00 (.01)
CEO tenure	.02 (.01)	.02 (.01)	.00 (.00)	.02 (.01)
COO presence	.26 (.22)	.25 (.22)	.02 (.03)	.02 (.03)
Prior performance[a, d]	.02 (.01)**	.02 (.01)**	.09 (.01)***	.09 (.01)***
Profitability (ROA)[a]	2.44 (.49)***	2.46 (.50)***	–.22 (.01)***	–.22 (.06)***
Sales growth[a]	2.85 (.61)***	2.91 (.62)***	—	—
CMO Presence[e]	.08 (.16)	.07 (.16)	.01 (.02)	.01 (.02)
Outsider CEO × CMO presence[f] (H_{9e})	—	–.26 (.36)	—	.04 (.05)

*$p < .10$
**$p <$ 05.
***$p < .01$.
[a] Tobin's q, sales growth, innovation, differentiation, prior performance, and profitability are the raw values less the median values at the two-digit SIC level.
[b] N = 156. Model 1a: Adjusted $R^2 = .46$, $F_{(15, 140)} = 9.73$, $p < .001$, Model 1b: Adjusted $R^2 = .46$, $F_{(16, 139)} = 9.13$, $p < .001$.
[c] N = 165. Model 2a: Adjusted $R^2 = .32$, $F_{(14, 150)} = 6.50$, $p <$ 001; Model 2b· Adjusted $R^2 = .32$, $F_{(15, 149)} =$ 6 12, $p < .001$.
[d] Prior performance is the value of the dependent variable in the year before the period of observation.
[e] We coded CMO presence as 1 (0) for firms with (without) a CMO for more than half the period of observation (2000–2004).
[f] Sample interaction between factor associated with CMO presence and CMO presence. All other similar interactions were also nonsignificant.
Notes Table shows parameter estimates along with standard estimates in parentheses. We averaged all time-varying variables across the period of observation (2000–2004).

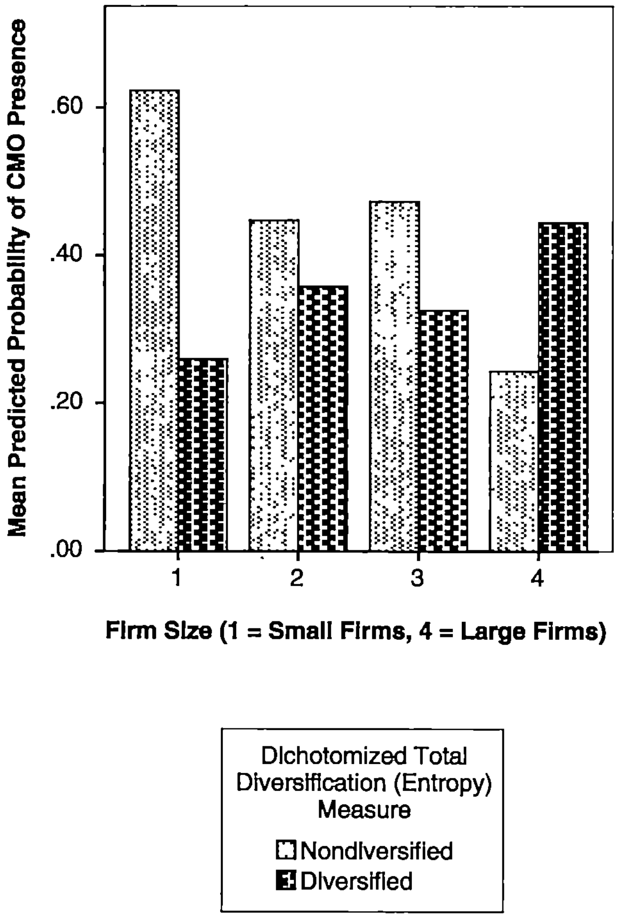
with the choice of having a CMO (for a summary of supported hypotheses, see Table 2). The presence of a CMO is related to innovation, differentiation, branding strategy, the CEOs operational experience in the firm, the diversification strategy and size of a firm, and, to some extent, the TMT's functional experience. However, CMO presence has no direct impact on firm performance, and having a CMO does not translate into higher levels of performance in the face of factors that are associated with CMO presence. In this section, we discuss implications for theory and practice, limitations, and suggestions for further research. We also present some illustrative examples of firms from our sample while discussing implications for theory.

Implications for Theory

To the best of our knowledge, we know of only one study that has identified factors related to CMO presence, as part of a conference board report on corporate marketing (Hopkins and Bailey 1984). Hopkins and Bailey (1984) studied a set of multibusiness firms and identified diversification as a key factor that explained CMO likelihood. Our study expands the scope of inquiry to include many more factors, it studies firms over time, and it updates the field on this phenomenon by examining data that are two decades from the aforementioned report. We identified the gaps addressed by our study as important to the field, given the efforts to understand marketing's role in the firm, especially at the level of corporate strategy. Notwithstanding other dimensions that capture marketing's importance to a firm, the CMO-related choice represents an important facet of marketing's influence in the TMT. This is evident in recent accounts of the CMO's role in the firm—namely, that of the customer advocate at the strategy table. For example, Joseph Tripodi, CMO of Allstate Insurance Company, defines his role as "the voice of the customer … particularly at the senior management table" (Crosby and Johnson 2005, p. 12). However, not all TMTs have CMOs, which leads to the question, Why do firms differ on this structural choice?

In exploring these phenomena, this research contributes to the dialogue on the role of marketing in the firm (Anderson 1982; Day 1992; Varadarajan 1992; Webster 1992; Wind and Robertson 1983; Workman, Homburg, and Gruner 1998), providing evidence of rationales for differences in marketing's influence at the corporate level of the firm. In the context of this research, these rationales helped identify specific strategic and structural factors, such as branding, innovation, diversification, and top executive experience, which are related to CMO presence. The rationale for the effect of the firm's branding strategy was that the criticality and complexity associated with building and protecting the corporate brand would make it more likely for firms with a corporate branding strategy to have a CMO. As an illustration of this rationale, consider the following two firms from our sample that are in the same industry—in

FIGURE 2
The Relationship Between Diversification and Probability of CMO Presence at Different Levels of Firm Size



this case, semiconductors: Advanced Micro Devices (AMD), which has a corporate branding strategy and sells all its product under the AMD brand name, and Micron Technology, which does not have a corporate branding strategy and sells products under the brand names Micron, Spectek, and Crucial. In line with the preceding rationale, we find that AMD has a CMO in all the years of observation, whereas Micron does not have a CMO in four of the five years. To the best of our knowledge, no prior research in the TMT literature has incorporated a firm's branding strategy in a framework of TMT's structural choices. Therefore, evidence of the effect of this variable marks a contribution of this research to organization theory.

The rationale of complexity in the marketing domain for the TMT also appears in the association between the strategic factors of innovation and differentiation and CMO presence. As firms pursue a strategy of innovation and differentiation, the need for a CMO seems to increase because a CMO is able to allay the TMT's resultant marketing-related uncertainty. Again, as an illustration of this rationale with respect to innovation, which we measure using the R&D-to-sales ratio, consider two firms in the same two-digit SIC 73: Oracle, whose average ratio between 2000 and 2004 was .12, and PeopleSoft, whose ratio was higher at .18. Similarly, with respect to differentiation, which we measure using the advertising-to-sales ratio, consider the following two firms that belong to the pharmaceutical industry: Abbott Laboratories, whose average ratio over the observation period was .01, and Bristol-Myers Squibb, whose ratio was higher at .08. In line with the rationales for these two factors, both Oracle and Abbott did not have a CMO in all the five years of observation, whereas PeopleSoft had a CMO throughout this period, and Bristol-Myers Squibb had one for four of the five years. Marketing complexity also helps explain why outsider CEOs are more likely to have CMOs. We expected this effect to be heightened for a new CEO, but the interaction between outsider CEO and the CEO's tenure as a CEO was not significant.[11]

The presence of a CMO is also explained by the firm's level of diversification, contingent on the size of the firm (see Figure 2). As firms diversify, they most likely adopt structures that reduce the TMT's complexity of managing diverse companies. As a result, a CMO is considered less essential in diversified firms. It is also likely that heads of the diversified businesses are reluctant to cede their power over decisions in the marketing domain to a CMO. However, this effect is more evident in firms that are relatively small in size. For example, if we compare Plantronics, an undiversified manufacturer of communication equipment, with Technitrol, a diversified manufacturer of electronic components, both with sales not exceeding $600 million, we find that Plantronics had a CMO, whereas Technitrol did not for the entire period of observation. Across relatively large firms, the effect of diversification seems to be reversed. For example, if we compare a large firm, such as Pfizer, which is undiversified, with another large firm, such as Procter & Gamble, which is highly diversified, we find that Pfizer did not have a CMO for four of the five years of observation, whereas Procter & Gamble had a CMO throughout this period. We suggest that this counterintuitive result is a result of size and diversity of businesses together confronting the TMT with increased complexity. As a result, the rationales of power and existing structures in small diversified firms are overcome in larger diversified firms. This counterintuitive component of our results differs from the finding of Hopkins and Bailey (1984), who find that unrelated diversification decreases CMO likelihood.

---

[11] We also checked other CEO-related variables, such as whether the CEO was also chairman, whether the CEO had been a COO of the firm, and whether the CEO's dominant experience was in output functions (i.e., marketing, manufacturing, and R&D) or in finance/law. None of these variables were significant. It is possible that the observation period of five years is a limitation in this regard because we observed a change in CEO in only 10.6% of the firm years in this period. Any CEO-related influence on the decision to have a CMO may be relevant only in the year the CEO is appointed.

We suggest that this could be due to the change in context in the past two decades and the result of deconglomeration, which has reduced the number of firms diversifying into unrelated areas (Varadarajan, Jayachandran, and White 2000).

The amounts of marketing and general management experience in the TMT were both found to be associated with CMO presence, though we did not find consistent support for these relationships. The rationale of homophily was supported, given the positive relationship found for marketing experience. It seems that marketing experience is required in the TMT for a CMO's role to be consistently appreciated. We also find evidence of the competing rationales of contingency theory and/or power dependence, given the negative relationship of general management experience to CMO presence. As we indicated in the discussion building up to this alternative hypothesis, it is not possible to determine whether only one or both of the logics of contingency theory and need for power are at play. Firms that take a contingency view consider general management experience in their top executives sufficient to address the tasks of a CMO. Alternatively, it might be the case that top executives with this experience are reluctant to let go of responsibilities in the marketing domain because these are bases for their power and influence within the TMT. To summarize, we find evidence of the rationales of (1) marketing-related complexity for the executives in the TMT, (2) criticality in the marketing domain, (3) structural arrangements within the TMT and the firm, (4) power dynamics within the TMT, and (5) homophily. Thus, further research in the domain of marketing's role in the firm and at the level of the TMT will be better informed given this evidence.

Implications for Practice

As firms evaluate their TMT compositions, we hope that this research aids them in making decisions in the context of CMO presence. Admittedly, the results of our research do not provide conclusive evidence of the benefits of a CMO, given the nonsignificant effects in the model of performance. Notably, however, CMO presence shared a significant, positive correlation with the Tobin's q measure of performance. Specifically, for the set of firms with/without a CMO across all years in the period of observation (N = 106), this correlation is .22 ($p < .05$). When this specification is relaxed to allow CMO presence to include firms with a CMO for more than half the period of observation (N = 134), this correlation is .21 ($p < .05$). However, these effects are dominated by other variables in the model of performance. Regardless, it is important to note that CMOs do not have a negative impact on performance.

However, firms can still benefit from the results of the model of CMO presence. Optimality of choices is implied or assumed in rationales that draw on the contingency view (Zeithaml, Varadarajan, and Zeithaml 1988). Under this assumption, CMO presence would be optimal for firms that aggressively pursue innovation and differentiation, for firms with corporate branding strategies, for firms with outsider CEOs, and for large diversified firms. Conversely, choices that are influenced by the rationale of power dependence may lead to suboptimal choices. Either of these two preceding rationales may be operating in the case of diversification in small firms. The power rationale implies that CMO absence may be suboptimal for small diversified firms, though the contingency logic based on existing structural arrangements implies that it may be optimal. Similarly, given these two rationales, it is not clear whether CMO absence is optimal when TMTs have high levels of general management experience. In this context, note that conceptual articles have argued that this experience is still not a substitute for an executive in the TMT who is explicitly focused on the marketing domain (Gilliatt and Cuming 1986). We hope that for these factors, the identification of the underlying rationales and their implied outcomes will still be of value to TMTs. The decision to have a CMO is complex, which is probably why fewer than 50% of the firms (in this and the Hyde, Landry, and Tipping [2004] studies) have such a role in place. In making top executives aware of the mechanisms that drive this choice in firms, this research has endeavored to take away the complexity that might be associated with it.

Limitations and Further Research

First, the nature of our sample does not allow for a systematic exploration of institutional effects, specifically those of imitation. A logic of institutional theory is that early adopters of a structural adaptation or innovation in an environment do so using contingency rationales and therefore are rewarded with performance gains; conversely, later adopters seek legitimacy and, in imitating other firms in their environment, do not achieve similar performance gains (DiMaggio and Powell 1983; Tolbert and Zucker 1983; Westphal, Gulati, and Shortell 1997). Therefore, in addition to providing evidence of imitation in CMO adoption, an exploration of institutional effects would shed light on the nonsignificance of the relationship between CMO presence and performance that we find in our study. However, to make any inferences about imitation, we must observe an entire population or, at the least, a substantially large set of firms that belong to a similar environment (in this case, the same industry) over a considerable period. The sample in this study is made up of firms drawn from multiple industries, and across the five-year period of observation in our study, we do not observe significant changes in CMO prevalence, with 66.5% of the firms consistently having/not having a CMO. Therefore, we are unable to comment on the presence or absence of institutional effects and suggest focusing on one industry over a longer period as an area for further research.

Second, our hypotheses for market concentration were not supported. A possible explanation might lie in a suggestion we made in the "Theory and Hypotheses" section; that is, whereas industries with low concentration have high marketing complexity, arguing in favor of CMO presence, marketing assets, such as customer satisfaction, may not be critical enough to warrant a CMO. We drew on Anderson, Fornell, and Mazvancheryl's (2004) finding of customer satisfaction's weakened impact on shareholder value in such industries. Possibly, this tension is also reflected in firms' CMO-related choices. For example, CMO presence was approximately 21% and 56% in SICs 28 and 73, respectively, which are both industries with relatively low values of concentration (HHIs of .02 and .04, respectively).[12] However, the lack of significant effects may also be due to the limited range of industries and concentration values (HHI values range from .02 to .18). Therefore, expanding the scope of this research to include additional industries might provide a clearer understanding of the role of market concentration. However, in analyses not shown here, we also found that several other environmental factors that are likely related to marketing complexity, such as industry sales growth, performance, and volatility, did not explain any additional variance in CMO presence at both two-digit and three-digit SIC levels. Again, one explanation may be that imitative forces are at play, in which case it would be beneficial to observe firms over a longer period of observation. Prior research has found that factors that influence structural adaptations may be strong or weak, depending on the stage of institutionalization of that adaptation (Tolbert and Zucker 1983; Zorn 2004). It might also be the case that the CMO position has been institutionalized in firms (Hannan and Freeman 1984); there is evidence of this, given the significant, positive correlations between CMO presence across the years of observation. Another explanation may be that the response to environmental marketing complexity is already captured in the firms' strategic and structural choices (Donaldson 2002).

Two other limitations in our research are related to the exclusion of firms with sales less than $250 million and firms that do not report R&D expenses. Given the lack of a significant correlation between size and CMO presence and the wide range of firm sizes that we included in our analysis, we suspect that the former is not a serious limitation and that our results should hold even for relatively smaller firms. We accept the latter as a limitation of this research and submit that our results cannot be generalized to firms in industries, such as retailing, that do not report R&D expenses.

We offer two final suggestions for further research. The first is to explore the effect of CMO presence on domains that are closely linked to the role of the CMO, such as brand equity. The second is to examine the characteristics of a successful CMO. There has been some concern in the field with respect to the phenomenon of CMO churn. Spencer Stuart, an executive recruiting firm, estimates the tenure of a CMO at two years; in comparison, a CEO's tenure is five years (Welch 2004). Reasons offered include a gap between the expectations and delivery from a CMO and the CMO's lack of strategic and/or line experience. Research that sheds light on these phenomena will be of value to top executives and CMOs.

---

[12] The presence of a CMO in SIC industries was 56% in SIC 73, 41% in SIC 35, 39% in SIC 36, 21% in SIC 28, 52% in SIC 38, 34% in SIC 30, 30% in SIC 34, 30% in SIC 25, 33% in SIC 26, and 33% in SIC 33.

REFERENCES

Aaker, David A. and Robert Jacobson (1994), "The Financial Information Content of Perceived Quality," *Journal of Marketing Research*, 31 (May), 191–201.

Anderson, Eugene W., Claes Fornell, and Sanal K. Mazvancheryl (2004), "Customer Satisfaction and Shareholder Value," *Journal of Marketing*, 68 (October), 172–85.

Anderson, Paul F. (1982), "Marketing, Strategic Planning and the Theory of the Firm," *Journal of Marketing*, 46 (Spring), 15–26.

Bantel, Karen A. and Susan E. Jackson (1989), "Top Management and Innovations in Banking: Does the Composition of the Top Team Make a Difference?" *Strategic Management Journal*, 10 (Special Issue), 107–124.

Bettis, Richard A. and Vijay Mahajan (1985), "Risk/Return Performance of Diversified Firms," *Management Science*, 31 (7), 785–99.

Calori, Roland, Gerry Johnson, and Philippe Sarnin (1994), "CEO's Cognitive Maps and the Scope of the Organization," *Strategic Management Journal*, 15 (6), 437–57.

Cannella, Albert A., Jr., and Michael Lubatkin (1993), "Succession as a Sociopolitical Process: Internal Impediments to Outsider Selection," *Academy of Management Journal*, 36 (August), 763–93.

Carpenter, Mason A., Marta A. Geletkanycz, and Wm. Gerard Sanders (2004), "Upper Echelons Research Revisited: Antecedents, Elements, and Consequences of Top Management Team Composition," *Journal of Management*, 30 (December), 749–78.

Chaganti, Rajeswararao and Rakesh Sambharya (1987), "Strategic Orientation and Characteristics of Upper Management," *Strategic Management Journal*, 8 (July–August), 393–401.

Chung, Kee H. and Stephen W. Pruitt (1995), "A Simple Approximation of Tobin's q," *Financial Management*, 23 (3), 70–79.

Crosby, Lawrence A. and Sheree L. Johnson (2005), "Change Agents: Chief Marketing Officers Are Positioned to Create Customer-Loyalty Centered Enterprises," *Marketing Management*, 12 (November–December), 12–13.

Cyert, Richard M. and James G. March (1963), *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice Hall.

Day, George S. (1992), "Marketing's Contribution to the Strategy Dialogue," *Journal of the Academy of Marketing Science*, 20 (Fall), 323–29.

——— (1994), "The Capabilities of Market-Driven Organizations," *Journal of Marketing*, 58 (October), 37–52.

——— and Prakash Nedungadi (1994), "Managerial Representations of Competitive Advantage," *Journal of Marketing*, 58 (April), 31–44.

Dearborn, DeWitt C. and Herbert A. Simon (1958), "Selective Perceptions: A Note on the Departmental Identification of Executives," *Sociometry*, 21 (2), 140–44.

DiMaggio, Paul J and Walter W. Powell (1983), "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review*, 48 (April), 147–60.

Donaldson, Lex (2002), *The Contingency Theory of Organizations*. Thousand Oaks, CA: Sage Publications.

Doz, Yves L. and C.K. Prahalad (1991), "Managing DMNCs: A Search for a New Paradigm," *Strategic Management Journal*, 12 (Special Issue), 145–64.

Finkelstein, Sydney (1992), "Power in Top Management Teams: Dimensions, Measurement, and Validation," *Academy of Management Journal*, 35 (August), 505–538.

——— and Donald C. Hambrick (1996), *Strategic Leadership: Top Executives and Their Effects on Organizations*. St. Paul, MN: West Publishing.

Fligstein, Neil (1987), "The Intraorganizational Power Struggle: Rise of Financial Personnel to Top Leadership in Large Corporations, 1919-1979," *American Sociological Review*, 52 (February), 44–58.

Gilliatt, Neal and Pamela Cuming (1986), "The Chief Marketing Officer: A Maverick Whose Time Has Come," *Business Horizons*, 29 (January–February), 41–48.

Hambrick, Donald C. (1981), "Environment, Strategy, and Power Within Top Management Teams," *Administrative Science Quarterly*, 26 (June), 253–76.

——— and Albert A. Cannella Jr. (2004), "CEOs Who Have COOs: Contingency Analysis of an Unexplored Structural Form," *Strategic Management Journal*, 25 (October), 959–79.

——— and Phyllis A. Mason (1984), "Upper Echelons: The Organization as a Reflection of Its Top Managers," *Academy of Management Review*, 9 (April), 193–206.

Hannan, Michael T. and John Freeman (1984), "Stuctural Inertia and Organizational Change," *American Sociological Review*, 49 (April), 149–64.

Hickson, David J., Christopher R. Hinings, C.A. Lee, Rodney E. Schneck, and J.M. Pennings (1971), "A Strategic Contingencies' Theory of Intraorganizational Power," *Administrative Science Quarterly*, 19 (June), 216–29.

Hise, Richard T. (1965), "Have Manufacturing Firms Adopted the Marketing Concept?" *Journal of Marketing*, 29 (July), 9–12.

Hitt, Michael A. and R. Duane Ireland (1985), "Corporate Distinctive Competencies, Strategy, Industry and Performance," *Strategic Management Journal*, 6 (July–September), 273–93.

——— and ——— (1986), "Relationships Among Corporate Level Distinctive Competencies, Diversification Strategy, Corporate Structure and Performance," *Journal of Management Studies*, 23 (July), 401–416.

Homburg, Christian, John P. Workman Jr., and Harley Krohmer (1999), "Marketing's Influence Within the Firm," *Journal of Marketing*, 63 (April), 1–17.

Hopkins, David S. and Earl L. Bailey (1984), *Organizing Corporate Marketing*. New York: The Conference Board.

Hyde, Paul, Edward Landry, and Andrew Tipping (2004), "Making the Perfect Marketer," *Strategy+Business*, 37 (Winter), 37–43.

Johnston, John (1991), *Econometric Methods*. New York: McGraw-Hill.

Keck, Sara L. and Michael L. Tushman (1993), "Environmental and Organizational Context and Executive Team Structure," *Academy of Management Journal*, 36 (December), 1314–44.

Kerin, Roger A. (2005), "Strategic Marketing and the CMO," in "Marketing Renaissance: Opportunities and Imperatives for Improving Marketing Thought, Practice, and Infrastructure," *Journal of Marketing*, 69 (October), 12–14.

Kohli, Ajay K. and Bernard J. Jaworski (1990), "Market Orientation: The Construct, Research Propositions, and Managerial Implications," *Journal of Marketing*, 54 (April), 1–18.

Kuester, Sabine, Christian Homburg, and Thomas S. Robertson (1999), "Retaliatory Behavior to New Product Entry," *Journal of Marketing*, 63 (October), 12–14.

Laforet, Sylvie and John Saunders (1994), "Managing Brand Portfolios: How the Leaders Do It," *Journal of Advertising Research*, 34 (September), 64–76.

Law, Sharmistha (2002), "Can Repeating a Brand Claim Lead to Memory Confusion? The Effect of Claim Similarity and Concurrent Repetition," *Journal of Marketing Research*, 39 (August), 366–78.

Lawrence, Paul R. and Jay W. Lorsch (1967), *Organizations and Environment*. Boston: Harvard University

Lazarsfeld, Paul F. and Robert K. Merton (1954), "Friendship as a Social Process: A Substantive and Methodological Analysis," in *Freedom and Control in Modern Society*, M. Berger, T. Abel, and C.H. Page, eds. New York: Van Nostrand, 18–66.

Mann, J. (1971), "Nominal and Effective Status of Chief Marketing Executives in Yorkshire Industry," in *Exploration in Marketing Thoughts*, G. Wills, ed. London: Crosby Lockwood, 268–76.

March, James G. and Herbert A. Simon (1958), *Organizations*. New York: John Wiley & Sons.

Markides, Costas (1995), *Diversification, Refocusing and Economic Performance*. Cambridge, MA: MIT Press.

McGovern, Gail J., David Court, John A. Quelch, and Blair Crawford (2004), "Bringing Customers into the Boardroom," *Harvard Business Review*, 82 (November), 70–80.

——— and John A. Quelch (2004), "The Fall and Rise of the CMO," *Strategy+Business*, 37 (Special Report), 1–8.

McPherson, J. Miller and Lynn Smith-Lovin (1987), "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups," *American Sociological Review*, 52 (June), 370–79.

Michel, John G. and Donald C. Hambrick (1992), "Diversification Posture and Top Management Characteristics," *Academy of Management Journal*, 35 (March), 9–37.

Mintzberg, Henry (1973), *The Nature of Managerial Work*. New York: Harper & Row.

Palepu, Krishna (1985), "Diversification Strategy, Profit Performance, and the Entropy Measure," *Strategic Management Journal*, 6 (July–September), 239–55.

Pfeffer, Jeffrey (1981), *Power in Organizations*. Marshfield, MA: Pittman Publishing.

Piercy, Nigel (1986), "The Role and Function of the Chief Marketing Executive and the Marketing Department," *Journal of Marketing Management*, 2 (3), 265–89.

Porter, Michael E. (1980), *Competitive Strategy* New York: The Free Press.

Rao, Vithala R., Manoj K. Agarwal, and Denise Dahlhoff (2004), "How Is Manifest Branding Strategy Related to the Intangible Value of a Corporation," *Journal of Marketing*, 68 (October), 126–41.

Ruekert, Robert W., Orville C. Walker Jr., and K.J. Roering (1985), "The Organization of Marketing Activities: A Contingency Theory of Structure and Performance," *Journal of Marketing*, 49 (Winter), 13–25.

Srivastava, Rajendra K., Tasadduq A. Shervani, and Liam Fahey (1998), "Market-Based Assets and Shareholder Value: A Framework for Analysis," *Journal of Marketing*, 62 (January), 2–18.

StataCorp (2005), *Stata Statistical Software: Release 9*. College Station, TX: StataCorp LP.

Taylor, Steven T. (2004), "Despite the Rejection of Some Major Law Firms, Chief Marketing Officers Are Here to Stay," *Of Counsel: The Legal Practice and Management Report*, 23 (November), 1–2, 16–17.

Thompson, James D. (1967), *Organizations in Action*. New York: McGraw-Hill.

Tolbert, Pamela S. and Lyrine G. Zucker (1983), "Institutional Sources of Change in the Formal Structure of Organizations: The Diffusion of Civil Service Reform, 1880-1935," *Administrative Science Quarterly*, 28 (March), 22–39

Varadarajan, P. Rajan (1992), "Marketing's Contribution to Strategy: The View from a Different Looking Glass," *Journal of the Academy of Marketing Science*, 20 (Fall), 335–43.

——— and Terry Clark (1994), "Delineating the Scope of Corporate and Marketing Strategy," *Journal of Business Research*, 31 (October–November), 93–105.

———, Satish Jayachandran, and J. Chris White (2000), "Strategic Interdependence in Organizations: Deconglomeration and

Marketing Strategy," *Journal of Marketing*, 65 (January), 15–28.

Vorhies, Douglas W. and Neil A. Morgan (2005), "Benchmarking Marketing Capabilities for Sustainable Competitive Advantage," *Journal of Marketing*, 69 (January), 80–94.

Webster, Frederick E., Jr. (1981), "Top Management's Concerns about Marketing: Issues for the 1980s," *Journal of Marketing*, 45 (Summer), 9–16.

——— (1992), "The Changing Role of Marketing in the Corporation," *Journal of Marketing*, 56 (October), 1–17.

———, Alan J. Malter, and Shankar Ganesan (2003), "Can Marketing Regain Its Seat at the Table?" Marketing Science Institute Report No. 03-003.

Weinzimmer, Laurence G., Edward U. Bond, Mark B. Houston, and Paul C. Nystrom (2003), "Relating Marketing Expertise on the Top Management Team and Strategic Aggressiveness to Financial Performance and Shareholder Value," *Journal of Strategic Marketing*, 11 (June), 133–59.

Welch, Greg (2004), "CMO Tenure: Slowing Down the Revolving Door," (accessed May 1, 2005), [available at http://content.spencerstuart.com/sswebsite/pdf/lib/CMO_brochureU1.pdf].

Westphal, James D., Ranjay Gulati, and Stephen M. Shortell (1999), "Customization or Conformity? An Institutional and Network Perspective on the Content and Consequences of TQM Adoption," *Administrative Science Quarterly*, 42 (June), 366–94.

Wiersema, Margarethe F. and Karen A. Bantel (1992), "Top Management Team Demography and Corporate Strategic Change," *Academy of Management Journal*, 35 (March), 91–21.

Wind, Yoram and Thomas S. Robertson (1983), "Marketing Strategy: New Directions for Theory and Research," *Journal of Marketing*, 47 (Spring), 12–25.

Woolridge, Jeffrey M. (2003), *Introductory Econometrics*. Mason OH: Thomson South-Western.

Workman, John P., Jr., Christian Homburg, and Kjell Gruner (1998), "Marketing Organization: An Integrative Framework of Dimensions and Determinants," *Journal of* Marketing, 62 (July), 21–41.

Zeger, Scott L. and Kung-Yee Liang (1986), "Longitudinal Data Analysis for Discrete and Continuous Outcomes," *Biometrics*, 42 (March), 121–30.

Zeithaml, Valarie A., P. Rajan Varadarajan, and Carl P. Zeithaml (1988), "The Contingency Approach: Its Foundations and Relevance to Theory Building and Research in Marketing," *European Journal of Marketing*, 22 (7), 37–64.

Zorn, Dirk M. (2004), "Here a Chief, There a Chief: The Rise of the CFO in the American Firm," *American Sociological Review*, 69 (June), 345–64.

David A. Schweidel, Peter S. Fader, & Eric T. Bradlow

Understanding Service Retention Within and Across Cohorts Using Limited Information

Service churn and retention rates remain central as constructs in marketing activities, such as valuation of service subscribers and resource allocation. Although extant approaches have been proposed to relate service churn to external factors, such as reported satisfaction, marketing-mix activities, and so on, managers often face situations in which the only information available is the duration for which subscribers have had service. In such cases, can they forecast service churn and understand the contributing factors, which may allow for subsequent intervention? The authors propose a framework to examine factors that may underlie service retention in a contractual setting. Specifically, they use a model of retention that accounts for (1) duration dependence, (2) promotional effects, (3) subscriber heterogeneity, (4) cross-cohort effects, and (5) calendar-time effects (e.g., seasonality). Then, they apply the framework to subscription databases of seven services offered by a telecommunications provider, mirroring the format commonly used to forecast future service churn (and to make managerial decisions). Across all seven services, the inclusion of promotional effects always improves the forecast accuracy of retention behavior, whereas including cross-cohort effects does not significantly improve it. In five of the services, customer heterogeneity, calendar-time effects, and duration dependence also contribute to improved forecasts. The authors use these results to understand how the expected value of a subscription differs across model specifications. They find considerable variation across model specifications, indicating that model misspecification can affect resource allocation decisions and other marketing efforts that are important to a firm.

Keywords: service retention, proportional hazards model, valuation

Retention remains a key construct for contractual service providers because it is essential for determining the value of existing and future subscriptions and for making resource allocation decisions. Modeling approaches have been proposed to understand churn patterns based on numerous factors, such as perceptions of quality (e.g., Zeithaml, Berry, and Parasuraman 1996), customer satisfaction (e.g., Bolton 1998), and the firm's marketing activities (e.g., Bolton, Kannan, and Bramlett 2000; Lewis 2005). Such research has furthered the overall understanding of the antecedents of service retention and the consequences of activities based on these factors. In practice, however, managers often face situations in which they have little or no external information available to them beyond the number of subscribers to a particular service. In these (realistic) situations, how accurately can service retention patterns be modeled (and forecast)?

For example, consider the case of Shaun, a hypothetical analyst working for a large telecommunications firm, who has only standard billing information available for a particular service offered by a cable television provider (e.g., high-speed Internet access). He knows the number of subscribers at any given time and the duration for which each of them has had service. He also knows about prior promotional programs the firm has run. However, the firm has not conducted surveys of its subscribers, so he does not have access to any information about customer characteristics or attitudinal measures. While recognizing that he cannot target individual subscribers with such limited information, can he understand the effects of different factors on retention and predict the number of subscribers who will discard service each month?

From his reading of the literature on retention forecasting, Shaun is aware of several factors that he can (and should) incorporate into his analysis. Prior research supports his initial observation that service churn decreases as a subscriber's tenure increases—that is, negative duration dependence (e.g., Hughes 2006; Reichheld 1996).

In his investigation, Shaun also notices that promotional offers (or, more generally, marketing activities) appear to affect the observed service churn (e.g., Lewis 2004). Some subscribers may leave after a promotional period ends, not because of increased dissatisfaction with the service but simply because of the resulting change in their price plan. Isolating subscribers' responsiveness to these offers can also enable the company to determine the best marketing strategy to help maximize the value of subscriptions (e.g., Lewis 2005).

David A. Schweidel is Assistant Professor of Marketing, University of Wisconsin–Madison (e-mail: dschweidel@bus.wisc.edu). Peter S. Fader is Frances and Pei-Yuan Chia Professor and Professor of Marketing (e-mail: faderp@wharton.upenn.edu), and Eric T. Bradlow is K.P. Chao Professor, Professor of Marketing, Statistics, and Education, and Academic Director of the Wharton Small Business Development Center (e-mail: ebradlow@wharton.upenn.edu), The Wharton School, University of Pennsylvania. The authors express their appreciation to the anonymous firm that graciously provided the data used herein.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (January 2008), 82–94

Shaun is also aware of the importance of accounting for unobserved differences across subscribers (i.e., subscriber heterogeneity) when modeling duration data (e.g., Morrison and Schmittlein 1980). Another concern he has is about differences that may exist between subscribers who started service last month and those who started last year. If the forecasting model has "staying power," he can use it to predict the service retention behavior of future subscribers (e.g., Neslin et al. 2006). However, he hesitates to apply the same forecasting model to the latest "cohort" of subscribers because there may be systematic differences in its behavior compared with older cohorts (i.e., cross-cohort effects). Another factor of interest that he has identified from prior literature is seasonality in retention patterns (e.g., Danaher 2002; Radas and Shugan 1998).

To complicate his work further, Shaun has been asked to analyze separately the churn patterns for multiple different services. Although he has explained that his ability to do so with only subscriber counts is limited, he is interested in knowing whether the same set of factors affects retention of each service. If so, he can apply the same forecasting model to each of the company's services; if not, he needs a flexible model that will let him understand which factors affect retention of which services. He decides to develop a general framework that he can easily apply to subscriber data from separate data sets to address this empirical question and to understand the drivers of service retention. This is the main objective of the current research.

As a starting point, consider the retention of Service A by four cohorts of subscribers that first signed up between February 2002 and May 2002.[1] These service retention curves appear in Figure 1, Panel A, and the empirical hazard rates appear in Figure 1, Panel B. Thus, Figure 1, Panel A, shows the proportion of initial subscribers from each cohort who still have service after t months, and Figure 1, Panel B, shows the proportion of remaining subscribers at month $t - 1$ who discontinue service at month t.

From these curves, a pattern emerges in service retention. Specifically, subscribers seem less likely to discard service within the first three months. After this period, the likelihood of customers to discard service suddenly increases, before ultimately decreasing in later months. In this research, we explore five specific factors, which may provide explanations for the observed service retention and hazard rate patterns both within and across cohorts (and for different services). These include (1) duration dependence, (2) time-varying marketing activity (e.g., promotional effects), (3) subscriber heterogeneity, (4) cross-cohort effects, and (5) calendar-time effects.[2] Because each of these factors may influence the hazard rate, we use a standard paradigm that allows for the incorporation of these factors—namely, the proportional hazard framework (e.g., Seetharaman and Chintagunta 2003). We now describe these factors briefly (a more formal technical description appears in the "Model Development" section).

Duration dependence allows the service churn rate for a subscriber to vary according to the length of time he or she has had service. For example, a decreasing churn rate as a function of time since acquisition might explain the observed decrease in overall service churn during the later months; however, this factor alone would not explain the increased service churn in the intermediate months in Figure 1, Panel A (and the increase in the hazard rate in Figure 1, Panel B). Thus, it is unlikely to be the sole contributing factor, but rather may be one of many.

Promotional effects, associated with a well-defined period of short-term marketing activity, can be seen clearly from the first portion of each cohort's retention curve. Throughout our data set, the service provider always offered a special three-month introductory period for new customers. Many service subscribers kept the service for the duration of the promotional period and then discarded it soon after the promotion ended, as reflected by the sudden increase in the empirical hazard rate in Figure 1, Panel B.

Although promotional activity might explain the initially low service churn and hazard rate, followed by the sudden increase after the promotional period, it is not enough to account for the decreased churn in the later months. This is where subscriber heterogeneity, the idea that service churn rates vary across subscribers, may come into play. Subscribers with high churn rates may drop service relatively quickly, leaving the firm with a smaller set of more stable subscribers in the long run and, thus, a curve that reflects greater retention over time. By simultaneously considering both duration dependence and heterogeneity, we can disentangle two competing effects: differences in service churn rates across subscribers and differences associated with tenure. Although both may lead to related retention patterns, they imply different stories and, consequently, different estimates of managerially relevant metrics, such as the value of a subscriber's subscription to a service, which we explore in greater detail subsequently.

Next, Panels A and B in Figure 1 also show some variability across the four cohorts, which may be attributable to systematic differences across cohorts. Although no obvious trend emerges graphically from Figure 1, Panels A and B, we need to allow for the possibility of such a cross-cohort effect when we examine a broader set of cohorts.

Finally, we consider calendar-time effects on service retention rate. In contrast to duration dependence, calendar-time effects focus on differences linked to the time of year rather than to the duration of a subscriber's service. Month-specific differences or seasonality can potentially influence the likelihood that subscribers will churn, regardless of their tenure, causing service churn to vary systematically for cohorts of all ages.

In summary, our objective is to build a general predictive model that can accommodate and sort out these competing explanations. By developing a predictive model, Shaun not only can forecast the future retention behavior of existing subscribers (in aggregate) but also can predict the retention patterns of future subscribers for which he currently has no information. The model itself is not revolu-

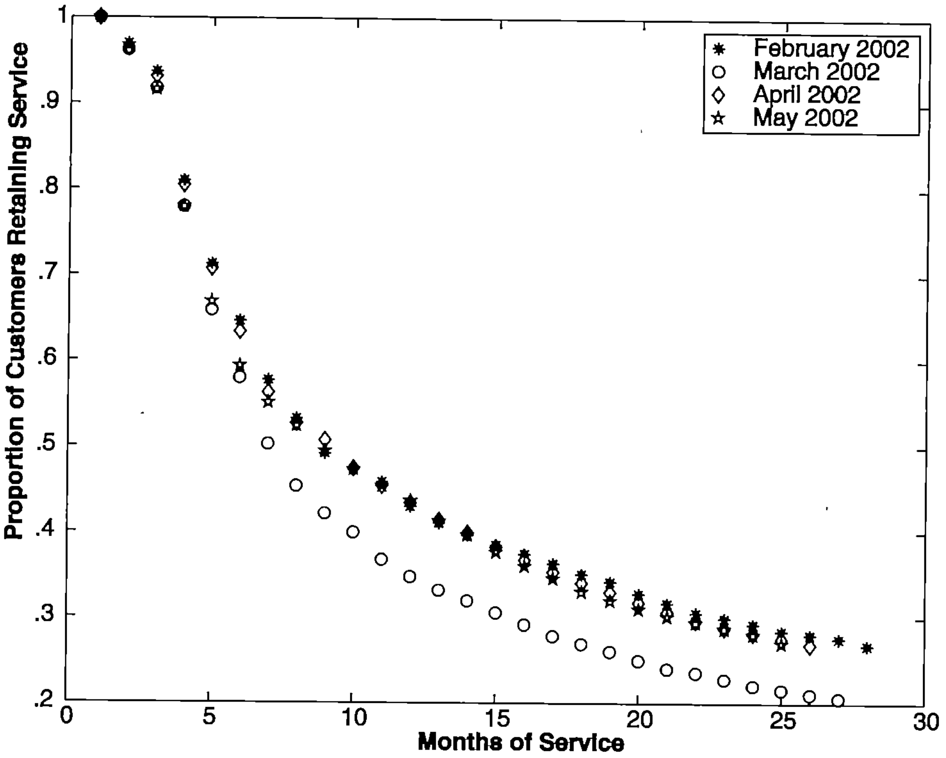
---

[1] The telecommunications company that provided the data preferred to remain anonymous; thus, we list services simply as A, B, C, and so on. We describe our data more fully throughout.

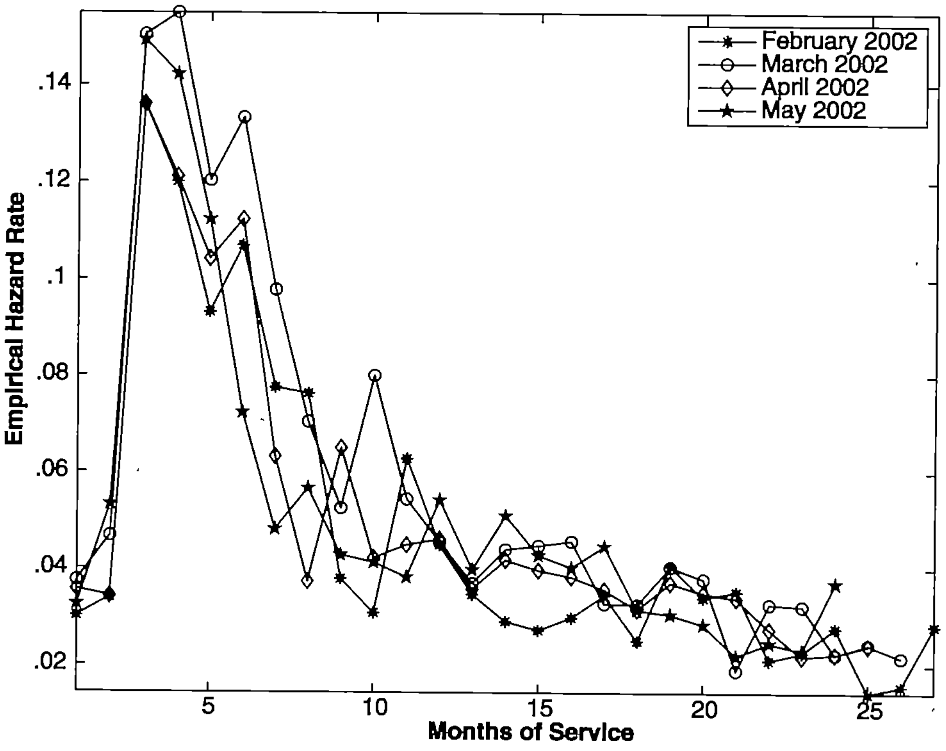
[2] Although these may not be the only possible effects, they represent the major factors described in the marketing literature, and our framework is flexible in its ability to accommodate other possible factors as well.

FIGURE 1
Cohort-Level Retention Behavior: A Selection of Cohorts

A: Cohort-Level Retention of Service A



B: Cohort-Level Empirical Hazard Rate of Service A



tionary; it closely resembles other duration models that have attempted to explain a varying set of drivers (e.g., Fader, Hardie, and Zeithammer 2003; Vanhuele et al. 1995). However, much of the extant research on service retention and churn relies heavily on external information at the level of the individual subscriber (including customer characteristics and attitudinal measures) to predict future subscriber behavior. In this research, we put forth a general modeling framework that can easily incorporate such information if it is available but can also forecast service retention of existing and future subscribers even when it is not available to the service provider, such as our hypothetical analyst, Shaun.

We carefully test these five factors, first for one service (as initially explored in Figure 1, Panels A and B) and then for a wider variety of services from the same provider. Across seven different services, we find strong support for the inclusion of promotional activity in modeling service retention. In addition, we find support for the inclusion of a combination of subscriber heterogeneity, duration dependence, and calendar-time effects. Notably, different services require a different set of components, highlighting the need for a general framework that can systematically examine competing explanations of behavior.

In the following section, we provide a review of the literature that has examined service retention. We then develop our modeling framework. Next, we describe our empirical analysis, in which we test various model specifications in a factorial design to understand systematically the effects of each of the five factors on customer retention. We then present detailed results of our empirical analysis for Service A and review the findings of our analyses for six other services offered. We conclude with a discussion of the implications and limitations of this research, as well as directions for future work.

Related Literature on Service Retention

In this section, we offer a brief review of some of the external factors that have been linked to service retention decisions, and we discuss potential limitations of these approaches for the context examined in this research. We then explore existing modeling approaches that can be used to model service retention solely on the basis of subscriber counts.

A considerable body of research has examined the link between satisfaction with a service and the duration for maintaining that service. Rust and Zahorik (1993) present a framework that links measures of satisfaction with retention and market share. Their framework allows a service provider to identify the dimensions of service satisfaction that will have the greatest impact on service retention and subsequently determine whether expenditures toward improving satisfaction are justified. Bolton (1998) finds that increased satisfaction increases the duration of service tenure. In addition, she finds that this effect is greater for subscribers who have had service for a longer duration. Her research also demonstrates the importance of satisfaction by assessing the increase in customer equity after service improvements are implemented. Apart from considering the effect of current satisfaction on subscribers' decisions to retain service, Lemon, White, and Winer (2002) incorporate expected future usage.

In addition to satisfaction, the link between service quality and service retention has received much attention in research. For example, Boulding and colleagues (1993) present a framework that links expectations of service quality to behavioral intentions. They hypothesize and find support for different effects of "should" and "will" expectations, which are combined with the service quality that actually occurred to develop an overall perception of service quality that affects future behaviors and expectations of quality. Rust, Zahorik, and Keiningham (1995) outline a methodology to estimate the return on investments in service quality, allowing providers to decide which improvement efforts should be undertaken. Zeithaml, Berry, and Parasuraman (1996) find that service quality affects subscribers' behavioral intentions with regard to the service, which in turn affect their retention decisions. They then link these decisions to financial outcomes. Bolton, Lemon, and Bramlett (2004) explore the impact of service renewal decisions in a business-to-business setting and find that subscribers are more likely to continue service after they experience a high level of service quality.

Other external factors that have been examined in relation to service retention include channel of acquisition (e.g., Reinartz, Thomas, and Kumar 2005), commitment and the effect of loyalty programs (e.g., Bolton, Kannan, and Bramlett 2000; Verhoef 2003), and payment equity (e.g., Bolton, Kannan, and Bramlett 2000; Bolton and Lemon 1999). In addition to these external factors, Keaveney (1995) identifies triggers that cause people to switch service providers. The most common type of trigger was a core service failure, such as billing errors or service mistakes. Among the other types of triggers were service encounter failures (e.g., speaking with an uncaring or unknowledgeable representative) and inconvenience to the subscriber.

When available, such attitudinal measures and external data can provide additional insight into the drivers of service retention decisions. However, as with the hypothetical Shaun, many service providers do not have easy access to such information. Instead, they often have little more than the number of subscribers from a cohort at a given time, because this can be directly extracted from internal billing information.

Fader and Hardie (2007) present a discrete-time probability mixture model that can be applied to the duration of service subscriptions. Although their modeling framework incorporates unobserved heterogeneity (as we do here), they assume a constant rate at which subscribers discontinue service. In addition, their framework does not accommodate cross-cohort effects or time-varying covariates identified previously in this research (or any of the other forms of external information that a provider could collect), the impact of which would be of interest to managers.

Note that extant research has also explored the link between defection in noncontractual (transactional) exchanges and customer value. In contrast to service retention, which we explore herein, the decision to defect in a noncontractual relationship is unobserved, and models that estimate customer value (e.g., Fader, Hardie, and Lee 2005; Schmittlein, Morrison, and Colombo 1987) rely on purchase histories, such as the time and number of prior purchases, which often do not have an analog in contractual exchanges. Unlike a noncontractual exchange, in which revenue is generated on purchase occasions until a customer defects, revenue in a contractual exchange is generated each period until a customer defects, which is observed. Thus, the models necessary to estimate retention (and, thus, value) in contractual and noncontractual exchanges fundamentally differ from each other.

As we described, we are not the first to consider service retention, contractual service retention, or its antecedents and consequences. However, our systematic exploration across multiple cohorts and services with limited information is unique—albeit that which is readily and commonly available.

Model Development

To forecast the number of subscribers retaining service in future periods, we propose a general specification using a parametric form for the "survival" probability, S(t) (i.e., the probability that a customer has maintained service until time t) and its complement, churn: $F(t) = 1 - S(t)$. That is, after calibrating the model on n periods of data, we can forecast the likelihood of maintaining service until time t by calculating S(t) for any $t > n$. This also enables us to estimate other related measures of interest to researchers and practitioners, such as the predicted number of customers who still have the service at time t $[N_0 \times S(t)]$, where N_0 is the number of customers who began service at time 0, and the number of customers expected to drop service between time t and t + 1 $\{N_0 \times [S(t) - S(t + 1)]\}$.

We focus on the hazard rate, the conditional rate of churn given that the customer has not already churned, which provides us with a well-established framework for duration models within which we can develop our general approach. We use a proportional hazards model, as is commonly used to account for various possible effects (e.g., Jain and Vilcassim 1991; Seetharaman and Chintagunta 2003), to incorporate the five factors listed previously. We do so within a mixture model specification to incorporate heterogeneity, in which the probability that a randomly selected customer i has not dropped service by time t is given by

$$S(t) = \int S[t|\theta_1, \boldsymbol{\beta}, \mathbf{X}(t)]g(\theta_1)d\theta_1, \tag{1}$$

where θ_1 is an individual-specific set of latent parameters, $\mathbf{X}(t)$ is a vector of covariates at time t, and $\boldsymbol{\beta}$ is the effect of these covariates. As such, Equation 1 may be considered a mixed-effects hazard model with both fixed and random components.

The mixture model in Equation 1 consists of two main components: $S[t|\theta_1, \boldsymbol{\beta}, \mathbf{X}(t)]$ and $g(\theta_1)$. The term $S[t|\theta_1, \boldsymbol{\beta}, \mathbf{X}(t)]$ specifies the probability that a customer maintains service until time t, which can be written in terms of its hazard function, $h[t|\theta_1, \boldsymbol{\beta}, \mathbf{X}(t)]$, which we specify to incorporate the five components of our model:

$$S[t|\theta_1, \mathbf{X}(t), \boldsymbol{\beta}] = e^{-\sum_{v=1}^{t}\left\{\int_{v-1}^{v} h[u|\theta_1, \boldsymbol{\beta}, \mathbf{X}(t)]du\right\}}. \tag{2}$$

Rather than assuming that all subscribers are homogeneous, the mixing distribution $g(\theta_1)$ allows for unobserved differences in subscribers' tendencies to discontinue service, because some may be inclined to do so after only a few months whereas others may be more reluctant. Next, we describe how each of the five factors that we consider (duration dependence, promotional activity, cross-cohort effects, calendar effects, and heterogeneity) is formulated within the mixture model presented in Equations 1 and 2.

Duration Dependence

As we noted previously, the likelihood that a subscriber will drop service may change according to the length of time for which he or she has had it. Therefore, we adopt the Weibull distribution for the baseline hazard, which is flexible and commonly used in proportional hazards models (e.g., Morrison and Schmittlein 1980; Seetharaman and Chintagunta 2003):

$$h_0(t|\lambda_1, c) = c\lambda_1 t^{c-1}. \tag{3}$$

The Weibull distribution nests the "strawman" exponential distribution when $c = 1$; if subscribers do not exhibit duration dependence in their likelihood to drop service, the model collapses to a constant hazard rate λ_1. Values of $c > 1$ yield an increasing hazard rate, implying that subscribers are more likely to discard service the longer they have had it. Conversely, $c < 1$ leads to a decreasing hazard rate, such that subscribers become less likely to discard service as their tenure increases. Depending on the nature of duration dependence, managers may want to allocate their marketing efforts toward "older" or "younger" subscribers.

Cross-Cohort Effects

The baseline hazard function given in Equation 3 is independent of the time at which subscribers began service (or the cohort to which they belong). To allow for systematic differences across cohorts, we incorporate a cohort-specific effect, $\ln[q(j)] = \beta_1(j - 1) + \beta_2(j - 1)^2 + \beta_3(j - 1)^3$, as a covariate affecting the baseline hazard function:

$$\begin{aligned} &h_0(t, j|\lambda_1, c, \beta_1, \beta_2, \beta_3) \\ &= c\lambda_1 t^{c-1} e^{\beta_1(j-1) + \beta_2(j-1)^2 + \beta_3(j-1)^3}, \\ &\text{for } j = 1, 2, 3, \ldots, \end{aligned} \tag{4}$$

where t is the time that has elapsed since customers began service in month j of the observation period. We employ a third-degree polynomial to allow for a range of possible patterns of cross-cohort effects, including a monotonically increasing or decreasing baseline hazard (as a function of the cohort j), a U-shaped (or inverse U-shaped) baseline hazard function, and a baseline hazard that increases, decreases, and then increases again. Note that for $j = 1$, the baseline hazard function in Equation 4 reduces to that given in Equation 3, implying that the cross-cohort effect can be interpreted as a scaling of the baseline hazard function relative to that of the first cohort:[3]

$$(5)\quad \frac{h(t, j|\lambda_i, c, \beta_1, \beta_2, \beta_3)}{h(t, 1|\lambda_i, c, \beta_1, \beta_2, \beta_3)} = e^{\beta_1(j-1) + \beta_2(j-1)^2 + \beta_3(j-1)^3},$$

$$\text{for } j = 1, 2, 3, \ldots.$$

This simple three-parameter model component, as given in Equation 4, should be sufficient to capture (or at least approximate adequately) any cross-cohort dynamics present in our data set.[4] If cross-cohort effects are present, certain cohorts of subscribers may be of greater value to the provider and thus targeted differently.

Promotional Activity

As we noted previously, we employ a proportional hazards model to incorporate time-varying covariates (e.g., promotional activity, calendar effects):

$$(6)\qquad h[t|\theta_i, \mathbf{X}(t), \boldsymbol{\beta}] = h_0(t|\theta_i)\exp[\boldsymbol{\beta}'\mathbf{X}(t)].$$

Our data provider (as previously described) indicated that a three-month promotional offer was standard practice. If promotional activity slows subscriber churn, as is expected, the hazard rate during the promotional period will be dampened compared with the nonpromotional period. Thus, when the promotional period ends, subscribers may become much more likely to discard service in a nonsmooth way, reflecting the higher hazard rate. To capture promotional activity, we define the variable Promo(t) such that Promo(t) = 1 for month t = 1, 2, 3, and Promo(t) = 0 otherwise. With an understanding of the impact of promotional activity, the service provider can determine whether the activity is actually worth its cost on the basis of the change in expected value with and without the promotion.

Calendar-Time Effects

Unlike duration dependence, which captures changes in service churn according to the length of time for which a customer has had service, calendar effects account for differences in the rate of service churn attributable to the time of year. For example, subscribers may be more inclined to discontinue certain services during the summer months if they will be on vacation. These effects will affect subscribers at different times in their service tenure, depending on the month in which they started each particular service. For example, subscribers beginning service in February will experience the "July effect" during their fifth month of service, whereas the June cohort will experience the same effect in its second month of service; thus, we can disentangle calendar time from cohort time, to which duration dependence and promotional activity are linked. In doing so, marketers can predict the parts of the year in which they are most likely to experience larger decreases in the number of subscribers and decide whether intervention is warranted.

Calendar-time effects are incorporated through month-specific shocks to the hazard function. Suppose that cohort j's tth month of service occurs in calendar month k (January = 1, February = 2, and so on). Let C(j, t) = γ_k, for k = 1, ..., 11, and $\gamma_{12} = 0$ (i.e., December is treated as a baseline). Combining promotional activity and calendar effects with the individual-level baseline hazard from Equation 4 yields the conditional probability of a subscriber from cohort j of a particular service maintaining it until time t, where $\Theta = \{c, \varphi, \tau, \beta, \gamma\}$ is the set of parameters common across individuals, given by

$$
\begin{aligned}
(7)\quad S[t, j|\lambda_i, \Theta, \text{Promo}(t)] &= \exp\Bigg[-\lambda_i\Bigg(\sum_{v=1}^{t}[v^c - (v-1)^c] \\
&\quad \times \exp\{\ln[q(j|\beta_1, \beta_2, \beta_3)] \\
&\quad + \beta_4\text{Promo}(v) + C(j, v)\}\Bigg)\Bigg] \\
&= \exp\{-\lambda_i \times A[t, j|\Theta, \text{Promo}(t)]\},
\end{aligned}
$$

where

$$
\begin{aligned}
(8)\qquad A[t, j|\Theta, \text{Promo}(t)] &= \sum_{v=1}^{t}[v^c - (v-1)^c] \\
&\exp\{\ln[q(j|\beta_1, \beta_2, \beta_3)] + \beta_4\text{Promo}(v) + C(j, v)\}.
\end{aligned}
$$

Subscriber Heterogeneity

To complete the proposed service retention model, we specify a mixture distribution, which accounts for unobserved subscriber heterogeneity in the likelihood of dropping a particular service. If all subscribers have the same propensities for discarding a service, the mixing distribution would simply be a spike [$g(\theta_i) = \lambda$], but this would be an extremely restrictive assumption. Instead, we allow for heterogeneity across service subscribers by assuming that each subscriber's λ_i is drawn from a gamma distribution:

$$(9)\qquad g(\lambda_i|r, \alpha) = \frac{\alpha^r \lambda_i^{r-1} e^{-\alpha\lambda_i}}{\Gamma(r)}.$$

We chose the gamma distribution not only for its flexibility but also because it is the conjugate prior for the Weibull distribution and therefore is commonly used as a mixing distribution for this purpose (e.g., Morrison and Schmittlein 1980). As such, the marginal probability of a subscriber in cohort j surviving until time t is found by integrating Equation 7 over the mixing distribution (Equation 9), as shown in detail in Equation 1, and is given by[5]

---

[3] The cross-cohort effect is incorporated in a manner consistent with the proportional hazard framework, where ln[q(j)] is treated as a stationary covariate. Note also that an intercept is not needed in the polynomial, because c serves in that role.

[4] Although we employ an "agnostic" third-degree polynomial in our empirical analysis, any functional form could be incorporated into the proportional hazard framework in a similar way.

[5] An acknowledged limitation of the proposed mixture modeling framework is that customers are assumed to have the same responsiveness to covariate effects (i.e., constant βs). In addition, other baseline hazard specifications could be chosen, though we use the Weibull to exploit its conjugacy with the gamma mixing distribution, allowing for a parsimonious model specification.

$$(10)\quad S[t, j|r, \alpha, \Theta, \text{Promo}(t)] = \left\{\frac{\alpha}{\alpha + A[t, j|\Theta, \text{Promo}(t)]}\right\}^r.$$

We have presented our complete modeling framework for cohort-level service retention. Note that the framework can easily accommodate behavioral measures, such as reported satisfaction and service usage as time-varying covariates, if they are available. These factors would be incorporated in a similar way as Promo(t) and C(j, t) into the hazard function. However, even when such detailed information is not available, we can still forecast the number of subscribers retaining service at the cohort level, as we demonstrate next.

Empirical Analysis

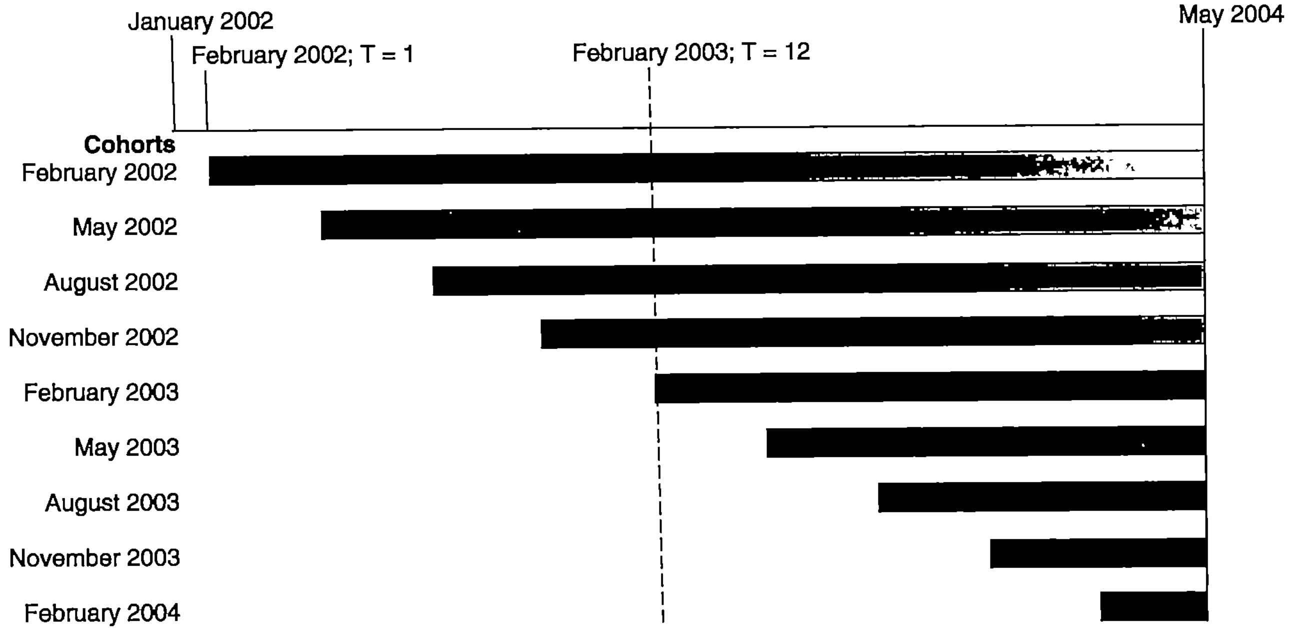
The data used in our empirical analysis were provided by a major telecommunications provider that offers a broad range of services to its customers. To evaluate the importance of the five factors, we separately fit a series of 32 models (defined by the full factorial of all combinations of the inclusion and absence of each factor) to seven services the company offers (which we denote as Services A–G). Some of them (Services A and C) are "base" services that do not require any of the others to operate and also require hardware, whereas the others are "add-on" services that create additional functionality. We begin by revisiting Service A (initially explored in Figure 1). Then, to assess the robustness of our findings, we apply the same set of models to subscription data for the other services.

A single regional billing center provided subscription information from January 2002 to May 2004, indicating (in aggregated monthly data) the number of households subscribing to each service at the end of each month. Thus, the first group that we observe from the time that it begins service is the February 2002 cohort. Figure 2 provides an illustrative example of the amount of data provided by each cohort; early cohorts are under observation for a longer period than later cohorts and consequently provide a larger number of observations for our analysis.[6]

The dashed line indicates the end of our chosen calibration period, at the end of February 2003 (T = 12 months). As Figure 2 illustrates, each cohort for calibrating the model is used for a different length of time. Although the cohorts that began service in or after February 2003 are not observed during the calibration period and therefore do not provide information to calibrate the model, they are still used to assess the (out-of-sample) fit of the model, thus providing a rigorous test of the model's forecasting ability. Specifically, for each cohort, we forecast the service retention numbers from March 2003 until the end of our data set in May 2004. We use the performance of the models in this out-of-sample forecast period to gauge the usefulness of the various model components.

However, the provided subscription information involves two forms of censoring that we must carefully

FIGURE 2
Diagram of the Cohort Analysis



---

[6] For ease of presentation, we show every third cohort in Figure 2. In our data, however, a new cohort comes under observation each month.

incorporate. First, the data are left censored because we only observe customers who maintain service for at least one month. To account for this, we calculate the probability of continued service conditional on having maintained service through the first month:

$$S^*(t, j) = P(T > t \mid t > 1) = \frac{S(t, j)}{S(1, j)}, \tag{11}$$

where S(t, j) is specified in Equation 10. Whereas the household data are left censored, the cohorts are not left censored; that is, we consider only cohorts that began service during the observation period. As such, we did not include households that began service before February 2002 because we did not observe when they started service.

Next, each observation is interval censored because we observe changes only in the number of subscribers for each service at the end of each month. Therefore, we construct a data set that contains the number of households from cohort j with service through time t, denoted as $N_{t,j}$. For each cohort j, households either maintain service through the calibration period T or discard service during the calibration period.

We consider households that maintained service throughout the entire calibration period. For a household starting service in month j (thus belonging to cohort j), keeping service through the calibration period implies that it maintained service for at least T – j + 1 months. Therefore, the probability that a household still has service at the end of the calibration period is given by $S^*(T - j + 1, j)$, and the number of households with service at the end of the calibration period from cohort j is $N_{T - j + 1, j}$.

The other possibility is that a household discards the service during the observation period. A household from cohort j could discard service in any interval (t, t + 1] for t = 1, 2, ..., T – j. Given the interval-censored nature of the data, the probability of a household from cohort j churning during the interval (t, t + 1] is given by $S^*(t, j) - S^*(t + 1, j)$, and the number of households discarding service during this interval is $N_{t,j} - N_{t+1,j}$.

From the households that maintain service and those that churn during the calibration period, we can construct the log-likelihood of the observed behavior of households from cohort j:

$$LC(r, \alpha, \Theta \mid j) = N_{T-j+1,j} \log[S^*(T-j+1, j)] + \sum_{t=1}^{T-j} \left\{(N_{t,j} - N_{t+1,j}) \log[S^*(t, j) - S^*(t+1, j)]\right\}, \tag{12}$$

where the first term accounts for households that maintain service through the entire calibration period and the second term accounts for households that churn during the calibration period, summing over all intervals in which the households could discard service. The log-likelihood for the full data set, consisting of all cohorts that began service during the calibration period, is as follows:

$$LL(r, \alpha, \Theta) = \sum_{j=1}^{T-1} LC(r, \alpha, \Theta \mid j). \tag{13}$$

Because the data are left censored, we include only cohorts that began service by T – 1 so that we observe at least one month of behavior from each cohort. We then obtained parameter estimates using maximum likelihood estimation.

As Figure 2 illustrates, we calibrated each of the 32 models using one year of data from the respective service, from the end of February 2002 through February 2003. To assess the overall forecasting ability of the models, we compare them on the basis of mean absolute percentage error (MAPE) for the out-of-sample period (from February 2003 to May 2004) and Bayesian information criterion (BIC) for the in-sample period.[7]

Results

Results for Service A

Analysis of the results from the 32 models fit to Service A revealed that the inclusion of heterogeneity always led to improved performance on the basis of both in-sample (based on BIC) and out-of-sample (based on MAPE) criteria, regardless of the other model components. For the models with heterogeneity, incorporating promotional activity also improved both in-sample and out-of-sample performance. Therefore, for ease of exposition, in Table 1, Panel A, we present the detailed results for the subset of 8 models in which duration dependence, cross-cohort effects, and calendar-related effects are varied but promotional activity and heterogeneity are always included. A comparison of the full series of 32 models appears in the Appendix. We first compare the relative performance of these 8 models and then discuss differences in parameter estimates and resultant inferences across the models in Table 1, Panel B.[8]

As we noted previously, the models incorporating heterogeneity and promotional activity perform significantly better than the corresponding models that omit these factors. Although the in-sample BIC suggests that the cross-cohort effect is necessary, the performance of the models in the holdout period indicates that it is overfitting the calibration data. In the case of Service A, we also find that duration dependence leads to systematic improvements in model performance (based on out-of-sample MAPE). Thus, the combination of unobserved heterogeneity, promotional activity, and duration dependence (Model 5) leads to the model with the lowest error in holdout churn forecasts, albeit not dramatically so over the model with just heterogeneity and promotional activity. We provide an illustration of the predictive ability of this model in Figure 3.

Note that incorporating the cross-cohort effects and/or the calendar-related effects leads to reduced accuracy in out-of-sample forecasting. Thus, after we account for promotional activity and subscriber heterogeneity, calendar-related and cross-cohort effects do not contribute to the

---

[7] The BIC is a commonly used penalization method that tries to prevent overfitting by penalizing the likelihood for each parameter added to the model. Lower values of BIC indicate better-fitting models. For more details on BIC, see Schwarz (1978).

[8] The small standard errors in Table 1, Panel B, are, in part, due to our large sample size. Small perturbations in the parameter estimates result in changes sufficiently large in the log-likelihood.

TABLE 1
Model Results for Service A

A: Model Performance for Service A

Model	Duration Dependence	Cross-Cohort	Calendar-Related	Log-Likelihood	BIC	MAPE
1				–45,946	45,961	10.8%
2			✓	–45,679	45,752	11.4%
3		✓		–45,927	45,959	19.5%
4		✓	✓	–45,624	45,714	46.8%
5	✓			–45,895	45,916	10.7%
6	✓		✓	–45,641	45,720	11.0%
7	✓	✓		–45,884	45,921	13.7%
8	✓	✓	✓	–45,612	45,706	46.1%

B: Model Parameters for Service A

Model	Duration Dependence	Cross-Cohort	Calendar-Related	(r, α)	$(\beta_1, \beta_2, \beta_3)$	β_4	c
1				(.68, 3.53)	—	−1.49	1
				(.01, .05)		(.02)	
2			✓	(.77, 6.43)	—	−1.44	1
				(.01, .08)		(.02)	
3		✓		(.66, 2.84)	(−.10, .02, −.00)	−1.48	1
				(.01, .04)	(.00, .00, .00)	(.02)	
4		✓	✓	(.34, .22)	(−1.01, .16, −.01)	−1.36	1
				(.00, .00)	(.00, .00, .00)	(.02)	
5	✓			(.32, 7.91)	—	−1.14	1.90
				(.00, .13)		(.03)	(.01)
6	✓		✓	(.37, 12.93)	—	−1.13	1.80
				(.00, .20)		(.02)	(.01)
7	✓	✓		(.33, 6.24)	(−.09, .01, −.00)	−1.16	1.83
				(.00, .10)	(.00, .00, .00)	(.02)	(.01)
8	✓	✓	✓	(.25, .28)	(−1.05, .16, −.01)	−1.19	1.43
				(.00, .01)	(.11, .02, .00)	(.02)	(.09)

forecasting ability of the model. Although heterogeneity across subscribers must be account for, there do not appear to be systematic differences across households based on the time at which they begin Service A. The way the cross-cohort and calendar-time effects are parameterized, as well as the data set, could warrant their inclusion in other studies, and thus our findings may not be entirely general in this regard. However, we have no reason to doubt the validity (or generalizability) of these observations, and we test for this using our six other services.

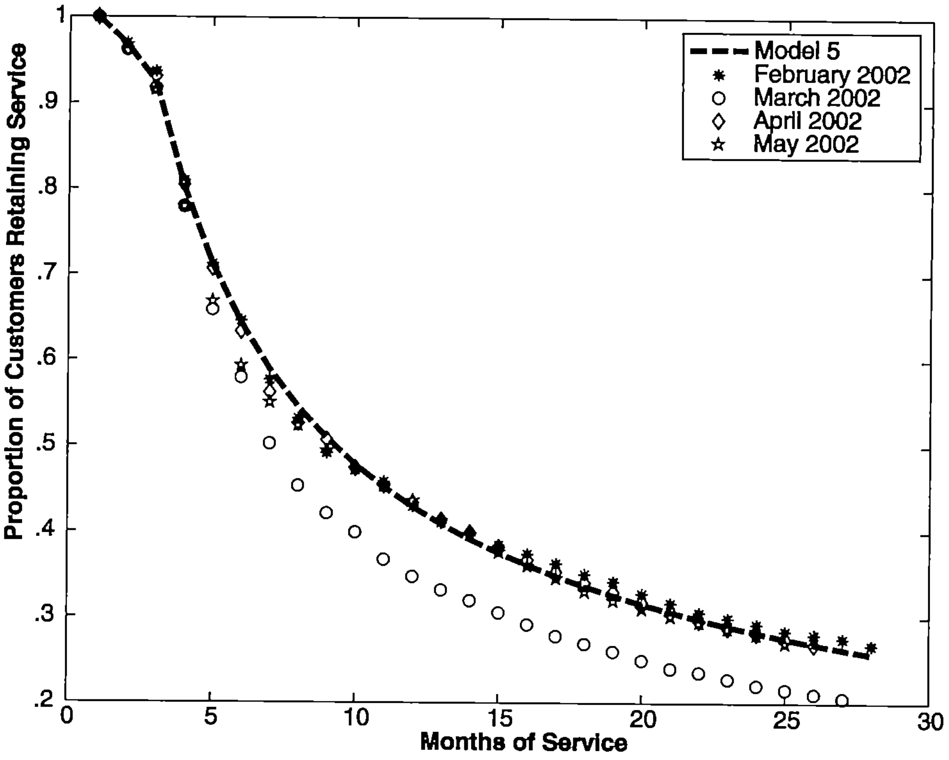
Next, we examine the parameter estimates derived from the models, which appear in Table 1, Panel B (along with standard errors in parentheses), to understand the managerial impact of using different models. First, we observe that there is positive duration dependence, as $\hat{c} > 1$ in Models 5–8, indicating that households are increasing in their probability of discarding service as their duration of service increases. This seems to run counter to previous research that has observed increasing retention rates over time (e.g., Reichheld 1996), as well as the empirical hazard rates that we observe in Figure 1, Panel B. As Follman and Goldberg (1988) note, however, omitting heterogeneity can lead to false conclusions about the effect of duration dependence, which is what we find here. Specifically, subscribers are not slowing down in the rate at which they drop service; rather, they are increasingly likely to drop service the longer they have subscribed to it, and the changes in the composition of remaining customers (those with higher retention rates maintain their subscriptions longer) best explain the observed increase in the aggregate retention rate. This finding emphasizes the need for a well-specified modeling framework because the type of duration dependence (positive or negative) can affect the way marketers focus their retention efforts on different groups of subscribers.

Second, with regard to the parameters governing the mixing distribution, the models that ignore duration dependence (Models 1–4) reflect a greater degree of homogeneity than the models with duration dependence (Models 5–8). (The coefficients of variation, the ratio of mean to standard deviation, are approximately 1.2 for Models 2–4 and 1.7 for Models 6–8.[9]) Thus, ignoring duration dependence may lead to the erroneous inference that subscribers behave more similarly in their tendency to discard Service A than is actually demonstrated by the data. This effect is not very

---

[9] Models 4 and 8 yield coefficients of variation that are greater in magnitude than those of Models 1–3 and Models 5–7, respectively, but the directional relationship holds because the coefficient of variation of Model 4 is less than that of Model 8.

FIGURE 3
Model 5 Fit for Service A



large, but it is systematic because it recurs for the other services as well. As Gupta, Lehmann, and Stuart (2004) demonstrate, changes in retention rates can lead to large changes in customer value. Thus, underestimating the impact of heterogeneity among subscribers' baseline service retention rates can lead to errors in calculating the expected value of a subscription.

The final finding of interest concerns the effect of promotional activity. The effect of promotional activity is stronger in Models 1–4 than in Models 5–8. Because promotional activity occurred during the first three months of service, to obtain a "clean" estimate of the effect of the promotion, we must also account for changes in the propensity to discard service that are attributable to duration dependence. Because $\hat{c} > 1$ for Models 5–8, subscribers are more likely to discard service as the length of time they have had service increases. Thus, by ignoring the effect of duration dependence, the effect of early promotional activity may be overestimated. If the provider were not to offer the promotion, under Model 1 (no duration dependence), 78.2% of subscribers would be expected to still have service after three months, a reduction of 15.6% compared with when the promotion is offered. Under Model 5, 82.9% of subscribers would be expected to remain, a reduction of 10.4%. By ignoring duration dependence in the baseline hazard function, managers may incorrectly attribute the reduction in churn solely to promotional activity, thus overestimating its effectiveness. Detecting such a difference can help managers avoid needless additional spending on promotional activities.

Generalized Findings for Other Services

Having found that, on the basis of out-of-sample analysis, the combination of promotional activity, subscriber heterogeneity, and duration dependence led to the best-performing model for Service A, we applied the same analytic procedure to the subscriber data for the remaining six services to assess the robustness of our findings. Table 2 summarizes the components of the "winning" models, again on the basis of out-of-sample MAPE performance for each service (though we also examined in-sample statistics, such as BIC).

We find the need for at least two of the highlighted factors for each of the seven services. For every service, we observe that promotional activity is always an important component of the best model, whereas the cross-cohort effect is not present in any of the winning models, despite its intuitive appeal. For most services, heterogeneity, duration dependence, and calendar-time effects also contribute significantly to model performance.[10] Across the winning models of different services, no more than two services

---

[10] In the models that incorporate heterogeneity, we find positive duration dependence ($\hat{c} > 1$); models that do not warrant the inclusion of heterogeneity demonstrate negative duration dependence ($\hat{c} < 1$).

TABLE 2
Summary of Winning Models

Service	Heterogeneity	Duration Dependence	Calendar-Related	Winning Model Value ($)	Minimum Value ($)	Maximum Value ($)
A	✓	✓		1,380	216	1,380
B	✓		✓	141	56	192
C		✓		2,376	856	9,472
D	✓	✓	✓	151	54	198
E	✓	✓	✓	143	53	203
F	✓		✓	150	51	208
G		✓	✓	74	60	290

share the same model specification, highlighting the need for a generalized modeling framework.

Notably, calendar-time effects are present in the winning models for all the add-on services the provider offers. Whereas Services A and C are base services that require hardware, Services B and D–G provide additional functionality that can be added or dropped by simply calling the service provider. In contrast, disconnecting a base service may require the customer to return hardware to the provider. A reasonable explanation posited by our data provider was that customers might disconnect these add-on services according to the time of year because of vacation patterns or the content of the service, which changed cyclically.

To highlight the managerial and economic importance of capturing the modeling components that affect service retention, we computed the expected value of a subscription, a measure akin to customer lifetime value, for each service (e.g., Berger and Nasr 1998). We can calculate this measure by multiplying the price less cost associated with the service by the expected duration for which the subscriber retains the service (e.g., Bolton 1998).[11] We perform this calculation for the eight models (see Table 2); promotional activity is always present, the cross-cohort effect is always omitted, and the inclusion of the remaining three factors is varied.[12] We present the expected value of a subscription under the winning model for the service, as well as the range of the expected value of a subscription under the eight model specifications. Note that these calculations could be performed using a full-factorial design with 32 models, but our primary interest is in assessing the combination of components present in the winning models.[13]

For the services in which heterogeneity is present in the winning model, the expected value of a subscription is closer to the high end of the range, whereas the converse is true for the services that do not require heterogeneity in the winning models. In general, omitting subscriber heterogeneity (when it is necessary) leads to a downward effect in estimates of subscription value, further emphasizing the need for a general, robust modeling framework for subscription duration.

In summary, we find that promotional activity, heterogeneity, duration dependence, and calendar effects are common elements in many of the models that yield the highest accuracy during the holdout period across the seven services considered. Furthermore, there are interactions among them; for example, omitting duration dependence appears to inflate the estimated effect of promotional activity and causes underestimation of the degree of heterogeneity across subscribers. Although detailed subscriber-level information, such as usage and satisfaction, would provide more insights into what affects subscribers' retention decisions, we can forecast the retention of current and future subscribers using limited cohort-level data.

Discussion and Conclusions

The motivation for this research was to develop a modeling framework that could forecast the number of subscribers who retain service with limited information and to understand the factors that contribute to retention behavior. Although much research has established the link among reported satisfaction, quality, and other antecedents, limited work has afforded such an understanding without this information. The flexible framework that we present is based on the extant modeling literature and can easily incorporate this external information when available. However, even without it, forecasts of the number of remaining subscribers from a given cohort are obtainable.

We examine five specific factors for their impact on retention: duration dependence, promotional activity, subscriber heterogeneity, cross-cohort effects, and calendar-time effects. Across seven different services, there is not a single set of factors that always leads to the best-performing model. Indeed, there is no overall model specification that wins for more than two of the seven data sets. On the one hand, the lack of a single winning specification may seem like a limitation of our analysis; on the other hand, the general framework is easily implemented (it takes less than a minute to run each model and can be done in widely available software, such as Microsoft Excel) and allows differences across services to be revealed. This systematic exploration is what managers need to do to understand their different sets of subscribers, because a "one-size-fits-all" model might take managers' attention away from the critical issues that we have covered. An area that warrants further consideration is understanding the characteristics of services that are affected by particular factors, such as calendar effects and duration dependence.

---

[11] Because price and cost information were not provided to us, for the purposes of our demonstration, we assume a cost of zero and prices that are consistent with the provider's current pricing scheme.

[12] Calculations for the full set of 32 models for each service are available on request.

[13] We compute the expected value for each of the 27 cohorts in our analysis and present the median result.

A further advantage of our formal modeling framework is its ability to aid managers in their evaluations of the effectiveness of marketing activity. As Table 2 shows, the model can provide estimates of the expected value of a service subscription. Managers can subsequently use this measure as a guide for determining how much they should be willing to spend on activities to induce a customer to subscribe to a particular service. In addition, by "turning off" the promotional activity (i.e., setting X[t] = 0 for all t), the same approach could be used to determine the expected increase in revenue from the specific promotion. Managers can use the calculation of expected value of a service subscription to tailor more effective marketing activities by changing features such as the time at which offers are available, the amount of the discount offered, or its duration (e.g., Lewis 2005). Through the proposed framework, by linking marketing activities to a financial metric (in this case, the expected value of a subscription), marketers can more effectively allocate their resources and make their marketing expenditures financially accountable (e.g., Rust, Zahorik, and Keiningham 1995).

Although this research focused on one service at a time, a key area for further research involves delving into the possible interplay among different products and services. Because many firms, such as our data provider, offer multiple contractual services, the development of integrated models for the adoption and retention of multiple services may assist managers in understanding multidimensional retention and churn issues, including cross-selling and subscriber valuation in a multiservice context. Unlike the model we present, such a framework may require individual-level subscription information rather than cohort-level information. Although such information could allow for individual-level marketing activities and more precise predictions of individuals' behavior, companies that have not implemented sophisticated systems can still benefit from using cohort-level data, as we demonstrate herein.

APPENDIX

Complete Model Results for Service A

Model	Heterogeneity	Duration Dependence	Promotional Activity	Cross-Cohort	Calendar-Related	Log-Likelihood	BIC	MAPE
1						–47,643	47,648	17.2%
2					✓	–47,220	47,283	17.4%
3				✓		–47,515	47,536	41.0%
4				✓	✓	–47,047	47,126	42.3%
5			✓			–46,419	46,430	24.2%
6			✓		✓	–46,015	46,084	26.6%
7			✓	✓		–46,400	46,427	30.6%
8			✓	✓	✓	–45,680	45,764	55.3%
9		✓				–47,327	47,338	27.3%
10		✓			✓	–46,877	46,945	29.8%
11		✓		✓		–47,277	47,304	43.4%
12		✓		✓	✓	–45,974	46,058	96.4%
13		✓	✓			–46,071	46,086	12.0%
14		✓	✓		✓	–45,784	45,857	13.1%
15		✓	✓	✓		–46,057	46,089	21.1%
16		✓	✓	✓	✓	–45,639	45,728	87.6%
17	✓					–47,643	47,653	17.2%
18	✓				✓	–47,220	47,288	17.4%
19	✓			✓		–47,515	47,541	41.5%
20	✓			✓	✓	–46,825	46,909	51.7%
21	✓		✓			–45,946	45,961	10.8%
22	✓		✓		✓	–45,679	45,752	11.4%
23	✓		✓	✓		–45,927	45,959	19.5%
24	✓		✓	✓	✓	–45,624	45,714	46.8%
25	✓	✓				–46,253	46,269	11.1%
26	✓	✓			✓	–45,971	46,044	11.9%
27	✓	✓		✓		–46,247	46,278	15.0%
28	✓	✓		✓	✓	–45,930	46,019	48.6%
29	✓	✓	✓			–45,895	45,916	10.7%
30	✓	✓	✓		✓	–45,641	45,720	11.0%
31	✓	✓	✓	✓		–45,884	45,921	13.7%
32	✓	✓	✓	✓	✓	–45,612	45,706	46.1%

REFERENCES

Berger, Paul D. and Nada I. Nasr (1998), "Customer Lifetime Value: Marketing Models and Applications," *Journal of Interactive Marketing*, 12 (1), 17–30.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17 (1), 45–65.

———, P.K. Kannan, and Matthew D. Bramlett (2000), "Implications of Loyalty Program Membership and Service Experiences for Customer Retention and Value," *Journal of the Academy of Marketing Science*, 28 (1), 95–108.

——— and Katherine N. Lemon (1999), "A Dynamic Model of Customers' Usage of Services: Usage as an Antecedent and Consequence of Satisfaction," *Journal of Marketing Research*, 36 (May), 171–86.

———, ———, and Matthew D. Bramlett (2004), "The Effect of Service Experiences Over Time on a Supplier's Retention of Business Customers," Marketing Science Institute Working Paper Series Report No. 04-119.

Boulding, William, Ajay Kalra, Richard Staelin, and Valarie Zeithaml (1993), "A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions," *Journal of Marketing Research*, 30 (February), 7–27.

Danaher, Peter (2002), "Optimal Pricing of New Subscription Services: Analysis of a Market Experiment," *Marketing Science*, 21 (2), 119–38.

Fader, Peter S. and Bruce G.S. Hardie (2007), "How to Project Customer Retention," *Journal of Interactive Marketing*, 21 (1), 76–90.

———, ———, and Ka Lok Lee (2005), "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42 (November), 415–30.

———, ———, and Robert Zeithammer (2003), "Forecasting New Product Trial in a Controlled Test Market Environment," *Journal of Forecasting*, 22 (5), 391–410.

Follman, Dean A. and Matthew S. Goldberg (1988), "Distinguishing Heterogeneity from Decreasing Hazard Rates," *Technometrics*, 30 (4), 389–96.

Gupta, Sunil, Donald R. Lehmann, and Jennifer Ames Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (February), 7–18.

Hughes, Arthur M. (2006), *Strategic Database Marketing*. New York: McGraw-Hill.

Jain, Dipak C. and Naufel J. Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10 (1), 1–23.

Keaveney, Susan M. (1995), "Customer Switching Behavior in Service Industries: An Exploratory Study," *Journal of Marketing*, 59 (April), 71–82.

Lemon, Katherine N., Tiffany Bartnett White, and Russell S. Winer (2002), "Dynamic Customer Relationship Management: Incorporating Future Considerations into the Service Retention Decision," *Journal of Marketing*, 66 (January), 1–14.

Lewis, Michael (2004), "The Influence of Loyalty Programs and Short-Term Promotions on Customer Retention," *Journal of Marketing Research*, 51 (August), 281–92.

——— (2005), "Incorporating Strategic Consumer Behavior into Customer Valuation," *Journal of Marketing*, 69 (October), 230–38.

Morrison, Donald G. and David C. Schmittlein (1980), "Jobs, Strikes, and Wars: Probability Models for Duration," *Organizational Behavior and Human Performance*, 25 (2), 224–51.

Neslin, Scott A., Sunil Gupta, Wagner Kamakura, Junxiang Lu, and Charlotte Mason (2006), "Defection Detection: Improving Predictive Accuracy of Customer Churn Models," *Journal of Marketing Research*, 43 (May), 204–211.

Radas, Sonja and Steven M. Shugan (1998), "Seasonal Marketing and Timing of New Product Introductions," *Journal of Marketing Research*, 35 (August), 296–315.

Reichheld, Frederick F. (1996), *The Loyalty Effect*. Boston: Harvard Business School Press.

Reinartz, Werner, Jacquelyn S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing*, 69 (January), 63–79.

Rust, Roland T. and Anthony J. Zahorik (1993), "Customer Satisfaction, Customer Retention, and Market Share," *Journal of Retailing*, 69 (2), 193–215.

———, ———, and Timothy L. Keiningham (1995), "Return on Quality (ROQ): Making Service Quality Financially Accountable," *Journal of Marketing*, 59 (April), 58–70.

Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987), "Counting Your Customers: Who Are They and What Will They Do Next?" *Management Science*, 33 (January), 1–24.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *Annals of Statistics*, 7 (2), 461–64.

Seetharaman, P.B. and Pradeep K. Chintagunta (2003), "The Proportional Hazard Model for Purchase Timing: A Comparison of Alternative Specifications," *Journal of Business and Economics Statistics*, 21 (3), 368–82.

Vanhuele, Marc, Marnik G. Dekimpe, Sunil Sharma, and Donald G. Morrison (1995), "Probability Models for Duration: The Data Don't Tell the Whole Story," *Organizational Behavior and Human Decision Processes*, 62 (1), 1–13.

Verhoef, Peter C. (2003), "Understanding the Effects of Customer Relationship Management Efforts on Customer Retention and Customer Share Development," *Journal of Marketing*, 67 (October), 30–45.

Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60 (April), 31–46.

Preyas S. Desai, Ajay Kalra, & B.P.S. Murthi

When Old Is Gold: The Role of Business Longevity in Risky Situations

The authors examine how a service firm's longevity may affect consumers' perceptions of risk in conducting business with the firm. Depending on the nature of information available to consumers, a firm's longevity can act in one of two ways to reduce the consumers' perceived risk: (1) Longevity may act as an extrinsic cue of the firm's quality, or (2) availability of a longer track record of an older firm may allow consumers to make more precise prediction of the firm's future quality. The authors develop a set of hypotheses that relate these two effects to consumers' preferences and test them in a series of four experiments. They find that a firm's age can act as an extrinsic cue of the firm's quality. When consumers also have information about firms' intrinsic attributes, the effect of a firm's age as an extrinsic cue is diminished under conditions of high involvement. The experiments also show that consumers prefer firms with longer track records, especially when firms' performance/quality levels are subject to more variability. However, when consumers have high aspiration levels, they prefer firms with shorter track records, even though the consumers perceive these firms as riskier choices.

Keywords: services, marketing strategy, advertising, risk, consumer behavior

Firms often advertise their year of inception; a casual examination of Yellow Pages advertisements for automobile repair, plumbers, and various other categories reveals that several firms highlight their starting year in their advertisements. For example, AXA Insurance's advertisements state that it has been "serving clients for over 137 years." This practice raises two questions: What are the advantages of communicating longevity in the business? and Why does this practice appear more prevalent in some industries (e.g., services) than in others? In other words, how does information about firm longevity influence customer behavior in some product categories?

In many purchase situations, consumers who are uncertain about the outcomes of their purchase decisions face one or more dimensions of perceived risk (Bettman 1973, 1975) and therefore engage in risk-reduction strategies (Mowen 1987), including searching for and processing more information (Dowling and Staelin 1994) or relying on extrinsic cues, such as price (e.g., Monroe and Dodds 1988), brand name (e.g., Dodds, Monroe, and Grewal 1991), advertising (e.g., Kirmani 1990), retailer reputation (e.g., Richardson, Dick, and Jain 1994), country of origin (e.g., Chao 1989), or warranty (e.g., Purohit and Srivastava 2001; Shimp and Bearden 1982).[1] We argue that business longevity may also act as a useful extrinsic cue about the quality of the firm and its products and thus reduce consumers' perceived risk. For example, consumers may believe that continued survival of a firm without offering strong benefits to consumers is unlikely or that a firm that has survived long enough will continue to survive in the future.

Compared with other extrinsic cues, firm longevity is unique in that it can be accompanied by information about the historical performance of a firm's intrinsic attributes (e.g., an airline's on-time record, a mutual fund's past returns), and older firms typically offer more of such information. Because, in general, consumers have more information about the past performance of older firms than of newer firms, this could affect their confidence in their assessment of the firm's performance. Even if the average performance of two firms is the same, additional data on an older firm enhance a customer's utility by reducing the uncertainty about future performance.

Because the nature of the information that a consumer possesses affects his or her perceptions of risk, we develop a conceptual framework that distinguishes among various purchase situations on the basis of the available information. In particular, we classify purchase situations into three groups: (1) when consumers have only the extrinsic cue of business longevity, (2) when consumers have both business

Preyas S. Desai is Professor of Business Administration, Fuqua School of Business, Duke University (e-mail: desai@mail.duke.edu). Ajay Kalra is Associate Professor of Marketing, Tepper School of Business, Carnegie Mellon University (e-mail: kalra@andrew.cmu.edu). B.P.S. Murthi is Associate Professor of Marketing, School of Management, University of Texas at Dallas (e-mail: murthi@utdallas.edu). The order of authorship is alphabetical. The authors thank Jim Bettman, Julie Edell, John Lynch, and Rick Staelin for their helpful suggestions.

---

[1] Schiffman and Kanuk (1994) list various strategies, including conducting more information search, being brand loyal, buying a well-known brand, buying from a reputable store, buying the most expensive brand, looking for warranties or money-back guarantees, and looking for independent lab results.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

longevity and current intrinsic attribute information available, and (3) when consumers have access to business longevity and historical information about intrinsic attributes. For each type of purchase situation, we develop hypotheses about the effects of a firm's longevity on consumers' perceptions of risk and choices. Then, we test the hypotheses in a series of laboratory experiments.

We find that customers prefer an older firm when they have no other information. When consumers also have access to information about intrinsic attributes, their choices are determined by the net effect of the intrinsic information and extrinsic cues of longevity. As consumers' involvement levels decrease, the relative effect of extrinsic cues strengthens. When historical performance data are available, we find that the firm with more available data is viewed as a less risky option and is preferred by consumers. This preference is stronger when consumers observe more variability in the performance data. Finally, we investigate the boundary conditions for the preference for older firms and find that when consumers' aspiration levels are high, they prefer riskier options and firms with shorter track records rather than longer track records.

In the next section, we develop the conceptual framework for the effects of business longevity on consumers' preferences. Subsequently, we present four experiments. Finally, we end with a discussion of the results and their implications.

Conceptual Framework

Our overarching objectives are to examine the effects of a firm's longevity on consumers' perceived risk in purchasing the product from the firm and then to relate the perceived risk to consumers' purchase decisions. In almost all purchase situations, consumers assess the expected quality of a product or service under consideration (Boulding, Kalra, and Staelin 1999; Boulding et al. 1993; Roberts and Urban 1986; Rust et al. 1999), but they rarely have all the necessary information to make a full assessment.[2] In service markets, the lack of information is more severe than in goods markets. Because of the large human component in service provision, service quality is inherently more variable over time than the quality of goods produced with modern plants and equipment (Zeithaml, Parasuraman, and Berry 1990). The human factor also makes the prediction of the expected quality of a service inherently more difficult.[3] The total risk due to the inherent variability is jointly determined by the level of inherent variability and the criticality of the decision for the consumer. For the same level of inherent variability in service quality, the consumer would perceive greater risk in situations in which his or her stake in the product is higher and the purchase decision is more critical for the consumer (Cox and Rich 1964).

As Dowling and Staelin (1994) note, additional information can reduce risk; we argue that firm longevity provides useful risk-reducing information. The effects of the firm's longevity on consumers' perceived risk depend on the amount of other intrinsic (i.e., product-/firm-related) information available to consumers. Because the amount of intrinsic information available to consumers can vary depending on the product category and situational factors, we envision three distinct possibilities: (1) situations in which consumers have little intrinsic information about various competitors, (2) situations in which consumers have some intrinsic information about the current attribute levels and/or service quality of various competitors, and (3) situations in which consumers have historical track record or performance data available about various competitors.

No Intrinsic Information Available

In some situations, consumers have little or no information about the quality of the service from a provider. For example, when they search for household repair services in the Yellow Pages, consumers gain little intrinsic information; if they have been in a traffic accident, they face a similar situation in attempting to find high-quality legal services. If the dates of inception of the repair company or the law firm were available to consumers, such information would provide a useful extrinsic cue of the service firms' level of experience and ability to survive in the marketplace.

Experience cues. Business longevity benefits a firm by ensuring its experience and progression through the learning curve, which in turn may enable it to provide services with less variance and possibly higher quality. Older firms can reduce their operating costs because of their learning and thus provide services more efficiently. They also are more likely to have a strong operational infrastructure, which is needed to provide reliable services (Berry and Parasuraman 1991; Blumberg 1991; Zeithaml, Parasuraman, and Berry 1990). Older firms have access to a wider window of opportunity to test and retest their service offerings and to reevaluate services after they are introduced, which also lowers service variance (Berry and Parasuraman 1991). At the individual employee level, well-trained, knowledgeable, experienced service employees are crucial (e.g., Anderson and Zemke 1998; Murray 1991), and because older firms are likely to have more established processes for training employees, their employees probably provide more consistent service. These factors all suggest that older firms can provide higher levels of service and can reduce variance in their service provision. As a result, consumers are likely to use this information to infer the firm's service quality and reliability.

Survival cues. In light of observed phenomenon that younger firms suffer higher mortality rates (e.g., Caves 1998), consumers may infer that a firm that has survived longer is more likely to survive in the future. Because consumers tend to want stable, durable relationships with their service providers (Crosby, Evans, and Cowles 1990), they may prefer older firms to newer firms, particularly in service categories in which the service is highly customized or service quality requires experience acquired over multiple encounters with customers (e.g., physicians, home repairs). In these cases, switching costs may be extremely high. Furthermore, information about a firm's business longevity may cause consumers to seek an attribution for its continued survival, and a logical attribution is that the firm provides high-quality service. Service providers commonly encourage this kind of attribution in their advertising messages, as in a local service provider's advertisement that reads, "We have been around for 81 years—we must be doing something right!"

---

[2] To simplify our discussion, we use the term "quality" to refer to all vector attributes of a product or service.

[3] However, as we discuss subsequently, the inherent variability of services can lead to other types of uncertainties and risks for the consumers.

In summary, when consumers have access to little intrinsic information, purchasing from an older firm seems to be a less risky choice, so consumers prefer an older firm to a newer firm. We expect that this effect is more pronounced when consumers have higher stakes in the purchase because they perceive more risk and therefore engage more in risk-reduction activities (Cox and Rich 1964; Dowling and Staelin 1994). This proposal is consistent with findings by Van Trijp, Hoyer, and Inman (1996), Erdem (1998), and Campbell and Goodstein (2001), who suggest that consumers prefer more established products (strong brand names) when facing greater risk.

> H_1: If no intrinsic information is available, consumers prefer firms with greater business longevity. This preference is greater for higher levels of perceived risk.

Intrinsic Information Available

Consumers often have access to information about a firm's business longevity and intrinsic attribute information, in which case they integrate both types of information to form quality assessments (e.g., Alba et al. 1994). When multiple cues exist, which one is the most salient depends on the relative diagnosticity or predictive value of the cues (Lynch, Marmorstein, and Weigold 1988; Slovic and Lichtenstein 1971). However, there is no generalizable finding about the diagnosticity level of cues, particularly when multiple cues are available and they do not provide consistent information.

Two recent articles have investigated this issue. Purohit and Srivastava (2001) classify attributes as either high scope, when their valence cannot be changed quickly (e.g., firm reputation), or low scope, when changes can occur faster (e.g., warranty). They argue that the high-scope cue moderates the diagnosticity of the low-scope cue. In addition, using negativity bias (Anderson 1965) as a theoretical basis, Miyazaki, Grewal, and Goodstein (2005) argue that when two extrinsic cues appear, the more negative cue is more salient, but extrinsic cues themselves are less informative when intrinsic cues are available and consumers are motivated to process information.

However, most prior research has been in the goods rather than services domain. Miyazaki, Grewal, and Goodstein (2005) argue that intrinsic cues should dominate extrinsic cues, but in services, intrinsic cues, such as level of assurance or empathy (e.g., Zeithaml, Parasuraman, and Berry 1990), are difficult to evaluate objectively and may be inferred on the basis of consumer expectations, formed from extrinsic cues (Boulding, Kalra, and Staelin 1999). Therefore, we argue that business longevity may imply important intrinsic attributes that are difficult to communicate in service settings (e.g., friendliness of staff). Even in the goods domain, Richardson, Dick, and Jain (1994) note that extrinsic cues, such as store versus national brand, affect perceptions of intrinsic cues. Whether business longevity is more diagnostic than an intrinsic cue remains an empirical question that depends on the abundance and concreteness of the available intrinsic attribute information. In the presence of intrinsic attribute information when consumers are more involved and, therefore, more likely to be motivated to process information (e.g., Petty and Cacioppo 1986), preference for business longevity is likely to decrease compared with when consumers are not highly involved.

> H_2: When current intrinsic information and extrinsic longevity cues are available, consumers use the longevity cue more in low-involvement conditions than in high-involvement conditions.

Quantitative Track Record Available

In markets such as financial services, consumers can access information about the past performance of service providers. For example, mutual funds must legally include their past performance records in their prospectuses, and such data are available in various financial media. Similar historical data exist for airlines and other service providers. Therefore, these records of past performance enable consumers to form specific expectations about future performance (Boulding, Kalra, and Staelin 1999; Rust et al. 1999).

When consumers have such information, the differences in the lengths of firms' track records may entail different levels of risk because the quality of consumers' quality estimates depends on the amount of information available. A longer track record leads to a more precise estimate of future quality. Research in Bayesian statistics and finance formalizes this idea into a construct called "estimation risk" (e.g., Bawa, Brown, and Klein 1979; Brown 1979; Coles and Lowenstein 1988; Klein and Bawa 1977; Stambaugh 1997), which represents the uncertainty (variance) involved in predicting expected quality on the basis of a firm's track record.

Subjective utility theory discusses similar notions, referred to as ambiguity and ambiguity aversion (for a review, see Camerer and Weber 1992). Ambiguity differs from risk in that risk refers to a situation in which the customer does not know for sure which of a possible set of outcomes will occur but can assign a probability to each outcome, whereas ambiguity means that the customer does not know for sure the probability distribution across outcomes. Ambiguity is due to missing information that could be known and is relevant. The similarity between estimation risk and ambiguity is best illustrated with an example that Camerer and Weber (1992) use. Consider a choice between placing a bet on two coins. The first coin has been tossed thousands of times with the conclusion that it is a fair coin. However, the second coin has been tossed only twice with the result of one head and one tail. Ambiguity regarding the true probability distribution of the second coin would suggest that placing the bets on the first coin is less risky because the risk of an incorrect belief about fairness of the coin has been lowered. This example clarifies the conceptual similarity between estimation risk and ambiguity. In the estimation risk domain, the amount of information varies, whereas ambiguity is broader, in that it also deals with a larger class of situations in which the customer has no information at all or data for a firm are missing.

According to the estimation risk literature (e.g., Bawa, Brown, and Klein 1979), the estimation risk for expected quality is formally captured by the variance term σ^2/N_i, where N_i is the number of observations available for a particular provider and σ^2 is the inherent variability in the service provider's quality.[4]

When consumers have access to track records of service providers, they will likely find a longer track record for an older firm than for a younger firm. Therefore, consumers will perceive a lower level of estimation risk for an older firm than for a younger firm, and all else being equal, they will prefer the former.

H_3: When consumers have access to service providers' track records, all else being equal, they prefer firms with longer track records.

H_4: When consumers have access to service providers' track records, all else being equal, they prefer firms with longer track records more when they face greater inherent variability than when they face less inherent variability.

Experiment 1: Business Longevity as an Extrinsic Cue

The objective of this experiment is to test H_1 by showing that consumers prefer older firms to newer firms when business longevity is the only differentiating information available and that their preference for older firms increases when perceived risk is greater. As we noted previously, the inherent variability of a service and the criticality of a purchase decision jointly determine the total risk due to variability for that purchase decision. It is difficult to change the inherent variability of a service without providing intrinsic information of the service, but in this experiment, our objective is to exclude the effects of intrinsic information. Therefore, we experimentally vary the total risk due to variability by manipulating the criticality of the decision.

Participants and Procedure

Participants (n = 136) were undergraduate business majors who participated for extra class credit. They were asked to imagine they were moving to a new city after graduation to take a new job. A description of their belongings was provided. Participants were told that after contacting several moving companies listed in the local Yellow Pages, they found that only two movers could provide services on their required moving date. They then viewed advertisements by the two hypothetical moving firms, Barrington and Lazzard. Immediately after viewing the advertisements, they provided measures of their purchase intentions and then responded to the manipulation checks and questions about their experience with moving companies.

We tested the hypotheses using a 2 × 2 × 2 × 2 between-subjects design. The manipulated factors were risk (low/high), business longevity of Lazzard (short/long), order in which the advertisements were viewed (Lazzard first/second), and text/picture version (two levels).

Stimuli

Brand names. In a pretest, participants considered 14 possible names and indicated how likely they were to select a moving company on the basis of the brand names given. We found no significant differences for the two selected brands.

Ad copy. We used full-page, black-and-white advertisements, in which the headline consisted of the brand name and the year the firm was established. The advertisement included a picture and information about the firm. We varied whether Barrington or Lazzard represented the older brand. To ensure that the effects were due to this treatment and not to the text of the advertisements, we also rotated two versions of the text and picture used. Thus, there were four advertisements for each brand.

Dependent variables. We measured purchase intentions with a 15-point scale anchored by "definitely select Lazzard/definitely select Barrington." Scores of 1–7 indicate an intention to purchase Lazzard, a score of 8 suggests indifference between the two brands, and scores of 9–15 imply intentions to purchase Barrington. Thus, the dependent measure captures both choice and relative preference intensity.

Independent Variables and Covariates

Risk. We manipulated risk by altering the description of the items the participants were told that they would be moving. In the low-risk situation, they were told that they owned few household items, including "some books, clothes, a bed, and some standard pieces of furniture acquired when you were a student." In the high-risk situation, participants also owned high-end electronic items, valuable china, antiques, and art pieces, several of which were described as fragile.

Firm age. To manipulate the age of the firm, we included the headline "Established: 1891" in the advertisement for the older firm. The year of establishment for the younger firm was 1996.

Experience with moving companies. We obtained a measure of participants' experience with moving companies through two nine-point scales ("novice/expert" and "not at all familiar/very familiar"). We used this as a covariate in the analysis.

---

[4] We thank an anonymous reviewer for pointing out that this expression for estimation risk is based on the assumption that all observations in the past track record are i.i.d. If the observations are correlated, each new observation provides less information than in the i.i.d. case. However, as long as the correlation is not perfect, each new observation offers additional information, and our qualitative arguments remain valid.

Manipulation checks. After completing the purchase intentions measure, participants responded to three nine-point measures of how risky they perceived the described situation ("not at all risky/extremely risky," "not at all concerned/extremely concerned," and "not at all worried/ extremely worried"). Finally, we measured perceptions of business longevity for both brands by asking participants to respond to two 11-point measures: "Compared to other moving companies, how many years do you think Lazzard/ Barrington have been in business?" and "Compared to other moving companies, when do you think Lazzard/Barrington started in business?"

Results

Manipulation checks. As Table 1 indicates, the manipulation of risk was successful. The main effect of risk was significant ($F_{1,130} = 20.22$, $p < .0001$), and there were no significant effects for business longevity or the business longevity × risk interaction (F values < 1). The main effect of perceived business longevity was significant when both Lazzard ($F_{1,131} = 278.27$, $p < .0001$) and Barrington ($F_{1,131} = 369.65$, $p < .0001$) served as the older firm. Neither the main effect of risk nor the business longevity × risk interaction was significant (F < 1).

We carefully pretested the two advertisements to ensure that they were equivalent with respect to the information provided and that no significant differences existed in attitudes toward the ads.[5] We measured attitude toward the ad using three seven-point scales ("very unfavorable/very favorable," "dislike very much/like very much," and "very bad/very good"). The manipulation check on attitude toward the ad revealed a significant, positive attitude toward the older firm (attitudes toward the Lazzard ad: $F_{1,135} = 59.52$, $p < .0001$; attitudes toward the Barrington ad: $F_{1,135} = 37.81$, $p < .0001$). However, there were no significant differences in the text or order effects, so it may be that participants evaluated the older firm's advertisement more positively because we measured attitude toward the ad after the main task. That is, the positive initial evaluation of the older brand may have led to favorable evaluations of the advertisement itself. Because participants rated the older firm's advertisements favorably for both brand and text versions, we conclude that any experimental effects are due to business longevity, not due to the ad executions.

Analysis. We examined the effect of advertisement order first. This factor did not affect responses to either of the two dependant measures (F values < 1). Next, we analyzed the effect of advertisement text and picture on both measures. The effects of variations of the ad text and picture on the purchase measure were also insignificant (F values < 1).

---

[5] All versions of the advertisements are equivalent, and manipulations of just business longevity within the same advertisement are not significantly different.

TABLE 1
Results of Experiment 1

	Low-Risk Condition		High-Risk Condition	
	Lazzard Older	Barrington Older	Lazzard Older	Barrington Older
Dependent Variable				
Purchase Intentions[a]	5.86 (3.28)	9.50 (2.60)	5.71 (2.99)	11.35 (2.25)
Manipulation Checks				
Risk perceptions	4.02 (1.76)	4.37 (1.70)	6.36 (1.52)	5.44 (2.10)
Perceived duration of Lazzard	7.55 (1.53)	2.77 (1.22)	7.39 (1.59)	3.41 (1.57)
Perceived duration of Barrington	2.97 (1.56)	7.36 (1.69)	2.67 (1.29)	7.10 (1.30)
Attitude toward Lazzard ad	4.99 (.99)	3.45 (.97)	5.07 (1.24)	3.70 (1.10)
Attitude toward Barrington ad	4.36 (.96)	5.27 (.64)	4.09 (1.25)	5.22 (.85)

[a] A response of 1 indicates that Lazzard is strongly preferred, 2–7 indicate progressively weaker preference for Lazzard, 8 indicates indifference between Lazzard and Barrington, and 9–15 show progressively stronger preferences for Barrington.

Therefore, we collapsed the cells into a 2 × 2 design. We conducted an analysis of variance (ANOVA) with category experience as a covariate to determine the treatment effects. For ease of exposition, we discuss the results for Barrington.

H_1 predicted that purchase intentions for the older firm would be greater than for the younger firm. The main effect of business longevity on purchase intentions was significant ($F_{1, 131} = 109.64$, $p < .0001$), providing strong evidence in support of H_1.

H_1 also predicted that the brand that has been longer in business would be more likely to be purchased in a high-risk situation than in a low-risk situation. The business longevity × risk interaction was significant ($F_{1, 131} = 8.91$, $p < .01$). When Barrington was the younger firm, purchase intentions decreased marginally in the high-risk situation compared with the low-risk situation ($F_{1, 131} = 2.57$, $p < .10$). When Barrington was the older firm, purchase intentions significantly increased when the risk increased ($F_{1, 131} = 6.01$, $p < .01$). These results support the notion that older firms are preferred when consumers perceive the risk as high.

Overall, Experiment 1 provides support for consumers' favorable evaluations of firms with greater business longevity. The results also show that preferences for greater longevity increase in high-risk situations. The results are consistent with prior literature that suggests that consumers prefer established brands in high-risk situations (Erdem 1998).[6]

Experiment 2: Business Longevity and Intrinsic Information

This experiment tests H_2 by examining the relative influence of current intrinsic attribute information and the extrinsic cue of business longevity on preferences.[7] We do this by providing both the date of inception and intrinsic information about two competing service providers under conditions of low and high involvement. The stimuli are constructed in such a way that the newer firm is better in terms of intrinsic information than the older firm. Thus, from H_2, we expect to observe a greater preference for the newer firm under high involvement than under low involvement.

Participants and Procedure

Undergraduate students at a major eastern U.S. university participated for a cash payment of $5. They evaluated two advertisements for dog kennels with the brand names Kennel Room and Petpourri. We used a 2 (low/high involvement) × 2 (business longevity short with strong intrinsic attributes/long with weak intrinsic attributes) × 2 (text/picture version) experimental design. In the high-involvement manipulation, participants were asked to assume that a relative who was visiting them next week accompanied by a dog had requested that they find a kennel in the local area. The pet was important to the relative, so it was important to find a kennel that would take good care of the dog. Participants were told that they should evaluate the advertisements for two kennels that were located conveniently. In the low-involvement manipulation, participants were asked to examine two advertisements for kennels and were told that the advertisements were adapted from firms in other states. Participants were instructed that they were part of a large group that was to proofread the advertisements for language and grammatical errors. We adapted these manipulations from instructions used previously in the literature (e.g., Sengupta, Goodstein, and Boninger 1997). After participants evaluated the advertisements, they indicated their purchase intentions and perceived quality of the two brands. Similar to Experiment 1, the 15-point purchase-intentions scale was anchored by "definitely select Petpourri" and "definitely select Kennel Room," and we measured perceived quality with two 15-point scales anchored by "Petpourri much higher quality/Kennel Room much higher quality" and "Petpourri much better/Kennel Room much better." The manipulation checks were performed at the end of the task.

Stimuli

The advertisements were half-page black-and-white, and both were on the same page. The headline for the older firm's advertisement stated, "Serving the area for 40 Years," whereas the year of establishment for the younger firm was 2003. On the basis of reviews of actual advertisements used by pet care centers, we selected the intrinsic attributes—pet exercise (dogs walked three times daily in two-acre exercise area versus separate exercise area), room heating (rooms heated/air conditioned versus in-floor heating), pet monitoring (caregivers there 24 hours a day versus continuous monitoring system), and veterinarian availability (veterinarians on staff versus veterinarians on call)—and we manipulated the strong versus weak intrinsic attribute condition for all four. Kennel Room represents the younger brand with strong intrinsic attributes in one condition and the older brand with weak intrinsic attributes in the other. We do not test all four combinations of strong and weak attributes and longevity, because an older firm with stronger intrinsic attributes should dominate a newer firm with weaker intrinsic attributes.

Results

Because there are no significant differences in either the brand name or the text/picture version conditions, we report the results of the pooled data.

Manipulation checks. We measured the strength of the intrinsic cues by asking participants to indicate which firm possessed the higher level of the four attributes on nine-point scales anchored by "Petpourri much better" and "Kennel Room much better"; then, we aggregated the responses pertaining to the four attributes. The effect of intrinsic cues was significant ($M_{Strong} = 5.21$ versus $M_{Weak} = 4.59$; $F_{1,80} = 6.44$, $p < .05$). To measure business longevity, we asked participants to indicate which brand had been in business longer on a nine-point scale anchored by "Petpourri much longer" and "Kennel Room much longer," and again, the effect of perceived business longevity was significant ($M_{New} = 2.32$ versus $M_{Old} = 7.88$; $F_{1,80} = 122.45$, $p < .001$). Following Sengupta, Goodstein, and Boninger (1997), we measured involvement according to how much effort participant reported they dedicated to evaluating the two brands, evaluating the information in the advertisement, or checking for grammatical and spelling errors (all on nine-point scales anchored by "no effort" and "tremendous effort"). The involvement manipulation was successful ($M_{High} = 5.58$ versus $M_{Low} = 4.96$; $F_{1,80} = 4.49$, $p < .05$).

---

[6] In Experiment 1, the hypothetical situation poses no real risk. Therefore, we replicate the experiment using monetary incentives and find the same choice pattern. For brevity, we do not report the results here.

[7] We thank an anonymous reviewer for suggesting this experiment.

Dependent variables. We coded the dependent measures as follows: 1–7 indicated a preference for the older firm, 8 indicated indifference, and 8–15 reflected a preference for the younger firm. H_2 predicted that high involvement would lead to stronger preference for the younger firm with strong intrinsic attributes than would low involvement. As we hypothesized, purchase intentions for the younger firm with strong intrinsic attributes increased significantly in the high-involvement condition ($M_{High} = 10.71$ versus $M_{Low} = 9.24$; $F_{1,80} = 3.39$, $p < .05$), as did its perceived quality ($M_{High} = 10.04$ versus $M_{Low} = 8.65$; $F_{1,79} = 4.26$, $p < .05$).

Note that in both involvement conditions, participants preferred the newer brand with strong intrinsic attributes. Two possible reasons may explain why the older brand with weak intrinsic attributes did not lead to being preferred in the low-involvement condition. First, the four intrinsic attributes together may have dominated the single business longevity cue. Second, the manipulation check reveals that even in the low-involvement manipulation, the absolute level of involvement remained moderate. Although these results provide evidence that the relative impact of business longevity increases in low-involvement situations, we suspect that the effect is much more significant when the intrinsic attribute information is weaker, because in such cases, the business longevity cue likely dominates intrinsic cues and results in greater preferences for older brands.

Experiment 3: Historical Track Record

The objective of this experiment is to test H_3 and H_4 by examining the impact of greater availability of historical data on consumers' preferences and their perceptions of estimation risk and total risk. We do this by providing historical track records of two competing firms about an intrinsic attribute such that even though both firms' records are comparable in terms of performance, one firm's record is longer than the other's.

Participants and Procedure

Participants were 122 business majors at a major eastern U.S. university who participated for extra course credit. They were asked to imagine that they were to fly to another city for a job interview and needed to select between two competing airlines that offered convenient schedules (Airline A and Airline B). Participants were informed that they would be given information about on-time performance for both airlines provided by a consumer magazine. They then saw the ratings for the two airlines, which indicated the percentage of time the airline arrived on time. Airline A began its operations in 1981, and Airline B began in 1999.

The 2 × 2 between-subjects experiment manipulated data variance (high/low) and data quantity (same/more). In the more-data condition, participants had access to performance ratings for all years the airline had been in business, so there were more data available for Airline A. In the same-data condition, participants were informed that the performance ratings were not available for the earlier years. Therefore, the performance ratings for the older airline (Airline A) were not provided for the first 18 years of its existence. As such, participants in this group had the same amount of data available for both airlines. Thus, they saw the data only from 1999.

Stimuli

We present the stimuli in Table 2. We designed the on-time arrival rate information so that the means, standard deviations, and ranges of any two data series would be identical. In all conditions, the mean arrival rate was 72.33%. In the low-variance condition, the standard deviation was 1.63,

TABLE 2
Experiment 1 Stimuli: On-Time Arrival Rates for Airlines

	Low Variance		High Variance	
Year	**Airline A**	**Airline B**	**Airline A**	**Airline B**
1981	72		80	
1982	74		61	
1983	73		52	
1984	72		97	
1985	74		52	
1986	70		73	
1987	72		54	
1988	70		90	
1989	74		54	
1990	70		97	
1991	73		61	
1992	74		98	
1993	72		87	
1994	74		54	
1995	74		56	
1996	70		95	
1997	74		52	
1998	70		89	
1999	70	74	52	84
2000	74	73	77	53
2001	74	71	99	69
2002	71	72	84	99
2003	72	70	53	52
2004	73	74	69	77

and the on-time arrival rates ranged between 70% and 74%. In the high-variance condition, the standard deviations for the same- and more-data conditions were 18.26 and 18.29, respectively, and the range of on-time arrival rates was between 52% and 99%. In the same-data conditions, we used the same set of numbers, so the means, variances, and ranges were identical. Our pretests indicated that participants used heuristics to form judgments of future performance, such as the number of times performance in one series exceeded that in the other series, most recent performance, and trends. We attempted to minimize these heuristics with our data, but because it is impossible to control all heuristics, we constructed the data to favor the younger firm, and thus we provide a more stringent test of our hypotheses.

Measures

Consumer preference. We measured preferences on a 15-point scale, ranging from "definitely Airline A" to "definitely Airline B," with an indifference point in the middle (8). Lower scores on this variable show participants' greater preference for Airline A.

Estimation risk. For each airline, we measured estimation risk by asking participants three questions: regardless of choice, based on past data, how comfortable they were in making their evaluations; how easy the evaluation decision was; and how certain they were that they made a good evaluation. The three nine-point scales were anchored by "not at all comfortable/extremely comfortable," "not at all easy/extremely easy," and "not at all certain/certain." An index formed by averaging the scores for these items resulted in Cronbach's alphas of .90 and .91 for Airline A and Airline B, respectively.

Overall risk. For each airline, participants responded to a nine-point scale that asked, "How risky do you think it will be to fly on Airline X?" ("not at all risky/extremely risky").

Results

Manipulation checks. At the end of the task, participants were asked about their perceptions of the data on two nine-point scales anchored by "low variance/high variance" and "low fluctuations/high fluctuations." A 2 × 2 ANOVA confirms that high-variance data ($M_{High} = 6.83$) were significantly different from low-variance data in both the same-data ($M_{Low} = 3.52$; $F_{1, 118} = 66.12$, $p < .001$) and the more-data ($M_{High} = 6.35$ versus $M_{Low} = 3.37$; $F_{1, 118} = 55.30$, $p < .001$) conditions. Participants also rated the following item on a nine-point scale: "Compared with Airline A, how much more information did you have about the arrival rates of Airline B" ("a lot less information/about the same amount of information [midpoint]/a lot more information"). Again, a 2 × 2 ANOVA confirms that participants believed that they had more information about the older airline in the more-data condition ($M_{More} = 2.58$) than in the same-data condition in the low-variance manipulation ($M_{Same} = 4.97$; $F_{1, 118} = 55.83$, $p < .001$) and the high-variance manipulation ($M_{High} = 2.68$ versus $M_{Low} = 5.17$; $F_{1, 118} = 79.32$, $p < .001$). Finally, we asked participants to rate the following: "Compared with Airline A, how long has Airline B been in business?" ("fewer years/same number of years [midpoint]/more years"). In all four conditions, participants scored Airline A significantly different from 5 (range 2.60–2.93) (all $t > 4.73$, $p < .001$).

Estimation risk. H_3 predicted that perceived estimation risk would be lower when the amount of data was more. For the older firm, a 2 × 2 ANOVA revealed significant main effects for data quantity ($F_{1, 118} = 11.14$, $p < .01$) and for data variance ($F_{1, 118} = 13.58$, $p < .001$). The interaction effect was not significant. In the high-variance condition, when more data were available, participants perceived estimation risk as significantly lower ($\bar{X} = 3.47$; $F_{1, 118} = 9.44$, $p < .01$) than when they had the same amount of data ($\bar{X} = 4.80$). In the low-variance condition, estimation risk was marginally lower in the more-data condition ($\bar{X} = 2.68$; $F_{1, 118} = 2.71$, $p < .10$) than in the same-data condition ($\bar{X} = 3.37$).

As we expected, when more data were available, estimation risk was relatively lower for the older firm than for the younger firm in both the high-variance condition ($F_{1, 118} = 4.06$, $p < .05$) and the low-variance condition ($F_{1, 118} = 5.76$, $p < .05$). With the same amount of data, estimation risk associated with older and younger firms did not differ significantly. That is, estimation risk depends on the amount of data rather than the cue of business longevity.

Risk. For the older firm, a 2 × 2 ANOVA reveals nonsignificant main effects for data quantity ($F_{1, 118} = 1.47$) but significant effects for data variance ($F_{1, 118} = 33.06$, $p < .001$). The interaction effect was not significant. As we predicted, in the high-variance condition, risk perceptions were significantly lower with more data ($\bar{X} = 4.55$) than with less data ($\bar{X} = 5.30$; $F_{1, 118} = 3.77$, $p < .05$). There were no significant differences in risk perceptions in the low-variance conditions.

As we also expected, in the more-data conditions, risk perceptions were relatively lower for the older firm in both the high- ($F_{1, 118} = 9.15$, $p < .001$) and the low- ($F_{1, 118} = 26.26$, $p < .001$) variance conditions. With the same amount of data, participants perceived marginally lower risk for the older firm when variance was high ($F_{1, 118} = 3.77$, $p < .10$) but no significant differences when variance was low.

Preferences. For the older firm, a 2 × 2 ANOVA reveals significant main effects for data variance ($F_{1, 118} = 3.85$, $p < .05$) and data quantity ($F_{1, 118} = 5.14$, $p < .05$) but no significant interaction effect. In H_4, we posited that preferences for the older firm, with its longer track record, would increase in high-variance conditions. This hypothesis is supported. In the high-variance condition, the preference intensity for the older firm in the more-data condition ($\bar{X} = 6.16$) was greater than that in the same-data condition ($\bar{X} = 7.53$; $F_{1, 118} = 6.81$, $p < .01$). In the more-data condition, preference intensity was greater for the older firm when variance was higher ($\bar{X} = 6.16$) than when variance was lower ($\bar{X} = 7.42$; $F_{1, 118} = 5.82$, $p < .01$).

Therefore, with the same historical performance—defined in terms of means, variance, and range—older firms benefit because they enjoy lower consumer perceptions of estimation risk. Furthermore, this benefit is enhanced when historical performances suggest higher variability. (The results of Experiment 3 appear in Table 3.)

Experiment 4: Preference for Shorter Track Records

Our preceding discussions assume that consumers are risk averse, which is valid in most purchase situations. However, are there conditions under which consumers would be risk seeking and choose a newer firm or a firm about which they have less historical data? Research on aspiration level (e.g., Crum, Laughhunn, and Payne 1981; Laughhunn, Payne, and Crum 1980; Payne, Laughhunn, and Crum 1980, 1981) points to a possibility along these lines. In our previous discussions, we considered situations in which consumers do not have prespecified aspiration levels. We now consider the effects of aspiration level to examine whether consumers can perceive lower risk in choosing a firm with longer history but still choose the firm with a shorter history.

As Simon (1955) and Payne, Laughhunn, and Crum (1980) note, consumers simplify choice problems by classifying outcomes as either satisfactory or unsatisfactory, depending on whether the outcome exceeds their aspiration level. Thus, consumers use a benchmark against which to evaluate outcomes. In a similar vein, Fishburn (1977) employs the notion of target return to define the risk of an alternative in a financial context. When managers are concerned with failing to meet their target return, they employ a convex utility function, consistent with risk-seeking behavior. However, if they are less sensitive to small failures, their utility function is concave, consistent with risk aversion. Kahneman and Tversky (1979) also show that consumers mentally encode outcomes as either gains or losses, depending on whether the outcome exceeds or falls below a reference point, and that the utility function is convex for losses. Crum, Laughhunn, and Payne (1981), Laughhunn, Payne, and Crum (1980), and Payne, Laughhunn, and Crum (1980, 1981) propose that outcomes are coded as losses or gains according to their relationships to the target or aspiration level. These authors find that people choose riskier alternatives when they face nonruinous losses but less risky alternatives when they face ruinous losses. Mittal, Ross, and Tsiros (2002) show that preferences for riskier options are moderated by the perceived self-capability of the decision maker and whether the decision is positively or negatively framed.

In our setting, the nature of risk differs significantly from that described in these other studies in that risk levels are based on the availability of data about two alternatives. In other words, two alternatives have the same distribution and expected payoffs but differ in the amount of data available to consumers. We expect that when people face losses, they display a preference for the riskier alternative—that is, the firm about which they have less information. Specifically, when consumers' target levels are higher than the expected performance of service providers, they believe that they will not meet their target and therefore will suffer a loss. In turn, they prefer the riskier alternative and choose the firm about which they have less information.

> H_5: All else being equal, consumers prefer firms with shorter track records when their target levels exceed firms' expected performance; otherwise, they prefer firms with longer track records.

We test H_5 in the context of an investment decision and thus delineate the boundary conditions for our theory about increased preference for firms with more data in riskier situations.

Participants and Procedure

We use a between-subjects design, in which the manipulated factor is a target investment return (high/low). Partici-

TABLE 3
Results of Experiment 3

	Low Variance		High Variance	
	Same Data	More Data	Same Data	More Data
Preference[a]	7.73 (1.26)	7.42 (2.08)	7.53 (1.91)	6.16 (2.69)
Estimation risk: older airline[b]	3.37 (1.46)	2.68 (1.70)	4.80 (2.03)	3.47 (1.50)
Estimation risk: younger airline	3.53 (1.35)	3.98 (1.54)	5.39 (2.17)	5.03 (1.55)
Overall risk: older airline[a]	3.30 (1.24)	3.39 (1.15)	5.30 (1.93)	4.55 (1.61)
Overall risk: younger airline	3.33 (1.55)	4.36 (1.33)	5.90 (1.84)	5.55 (2.14)

[a] Lower numbers indicate higher preferences for older firms.
[b] Higher numbers indicate more inherent variability

pants were asked to imagine that they were planning to invest $10,000 in a mutual fund. The stimuli were similar to those used in Experiment 3; participants selected between two mutual funds, A and B. Mutual Fund A was established in 1977, and Mutual Fund B started in 1993. We provided the annual percentage return of each fund for each year the firm was in existence. The average return of each fund was 8.81% with a variance of approximately 2.42%, and the range of returns for both firms fell between 5.1% and 12.1%. In the high-target condition, participants were asked to assume that after determining their consumption needs, they needed to get an average annual percentage return of 12% in the next five years; in the low-target condition, the target return required was 8%. In the 24-year history of the older firm, it had achieved the target of 12% three times; in the 8-year history of the younger firm, it had reached the target once. Participants were asked to examine the returns for both firms and to indicate the mutual fund they would select. We obtained measures of relative estimation risk by asking participants, "If you have to make a forecast of returns for the next five years, which fund will you be more confident in your forecast." Similarly, we measured expected variability by asking participants, "Which fund's returns do you think will be more volatile?" Finally, we asked participants to respond to a series of questions to explain their choices. The key question asked about their agreement with the statement, "Given the same past performance by both funds, the newer firm has a better upside potential in the future than the old firm" (1 = "strongly disagree," and 7 = "strongly agree").

Results

Of the participants, 75.9% (n = 29) selected the fund with greater longevity when the target return was low. The preference for the older fund decreased when the target return was higher, with only 35.5% (n = 31) selecting the fund with greater longevity. The difference was highly significant. As we expected, the high-target condition led to a significant decrease in preference for the older fund: $\chi^2(1) = 10.2$, $p < .001$. Notably, 64.5% of the participants selected the low-longevity firm in the high-target condition, even though they recognized that it entailed greater estimation risk (64.5% of the participants). Participants also stated that the newer firm's performance was more likely to be volatile (74.2% of the participants). Finally, the mean agreement response to the statement on the relatively greater upside potential of the newer firm was significantly higher in the high-target condition ($\bar{X}_{\text{High Target}} = 3.94$) than in the low-target condition ($\bar{X}_{\text{Low Target}} = 2.76$; $F_{1,58} = 19.05$, $p < .0001$). The results support the notion that if two alternatives have similar past performance, high aspiration levels may lead to preference for the alternative that has greater estimation risk. In particular, this behavior would be observed if the aspiration levels were close to or higher than the upper bounds of the past performance levels of low-estimation-risk alternatives. Although this experiment provided longitudinal data on performance, the results are likely to hold even in cases of more limited information, such as simple averages of previous ten years' performances.

Conclusions

In many service markets, a large human component is required in delivering the service to consumers. This makes service quality inherently more variable and imposes greater uncertainty on the consumers. We identify two ways that longevity of a firm can reduce consumers' perceived risk: (1) The longevity of a firm can act as an extrinsic cue about the firm's quality, and (2) an older firm may naturally provide a longer track record about its past performance, which allows consumers to predict the firm's future quality with greater precision. A summary of the results appears in Table 4. Experiment 1 shows that when only information about business longevity is available, business longevity acts as an extrinsic cue, increasing consumer preferences for the older firm, particularly in high-risk situations. This result is consistent with previous findings that greater risk leads to more cue usage, such as relying on established brand names (Erdem 1998), favoring standard product appearances (Campbell and Goodstein 2001), and decreasing variety seeking (Van Trijp, Hoyer, and Inman 1996). When intrinsic attribute information is also available, con-

TABLE 4
Summary of Results

Available Information	Effects
Only business longevity information is available; no intrinsic information and no track record are available.	•Consumers prefer firms with greater business longevity. •Preference is stronger at higher levels for perceived risk.
Longevity and qualitative intrinsic information is available; no track record is available.	•Consumers rely more on the longevity information under low involvement than under high involvement.
Quantitative track record is available; no qualitative intrinsic information is available.	•Consumers prefer firms with a longer track record (usually, firms with greater longevity) when their targets are lower than expected firm performance; otherwise, they prefer firms with a shorter track record (usually, firms with lesser longevity). •Consumer preference for firms with a longer track record (usually, firms with greater longevity) increases as the inherent variability increases.

sumers integrate information about both cues, but in low-involvement situations, in which consumers lack motivation to process information, the diagnostic value of the business longevity cue increases.

When business longevity information is supplemented with historical performance information, additional data reduce the level of estimation risk, regardless of whether the data indicate high or low variance. However, consumers' perceived risk falls only when additional data are available in high-variance environments, in which case the data also lead to increased preference for the older firm.

We also identify a boundary condition for preference for older firms. Consumers with high aspiration levels tend to prefer newer firms despite viewing them as riskier. This is apparently because they view newer firms as having greater upside potential than older firms with similar performance levels.

Implications

Our results provide several managerial implications in the areas of segmentation, targeting, and marketing communications. Overall, our results help managers in the service sector understand how the longevity of their firms can affect consumers' perceptions of risks involved in dealing with them. In addition, the results of Experiment 2 show that the relative impact of longevity as an extrinsic cue diminishes when consumers are more involved. With this knowledge, managers can design appropriate messages that can either build on positive aspects of their longevity or address concerns that arise from their firms' relative youth. These messages can be further shaped to place greater or lesser emphasis on the age of the firm relative to the intrinsic attribute information, depending on the level of consumers' involvement in the given product category. If there are significant differences among consumers in the level of their involvement, firms should explore the possibility of segmenting the market in terms of the consumers' involvement. Older firms may enjoy competitive advantage over younger firms with similar intrinsic attributes among lower-involvement consumers. Conversely, younger firms may want to target more involved consumers, especially if they have superior intrinsic attributes.

We also demonstrate the importance of providing additional historical track record information to consumers. Providing more information to consumers about past performance can enable them to develop more precise estimates of a firm's future quality and to perceive lower risk in dealing with the firm. This risk reduction helps a firm when consumers have lower aspiration levels but hurts the firm when consumers have higher aspiration levels. Research already has established that the frequency of information affects consumer perceptions and choice (e.g., Alba and Marmorstein 1987), and we demonstrate that the quantity of data exerts a similar impact, even when the data are similar in terms of mean and variance. Therefore, these results have managerial implications for both long-standing and relatively younger firms. Firms that have been in business for a long time should communicate that information to reduce consumers' perceived risk when consumers have relatively lower aspiration levels. Among firms with similar business longevity, those that provide more information about their track records will be regarded as the less risky choice and thus will be preferred by consumers. In contrast, younger firms need to be aware of their relative disadvantage and should develop strategies to compensate for it. Such firms should highlight other risk-reduction mechanisms, such as performance guarantees, warranties, and endorsements from reputable third parties. When there is a significant amount of heterogeneity among consumers in a market in terms of their aspiration levels, relatively younger firms should target higher-aspiration-level consumers, and relatively older firms should target lower-aspiration-level consumers.

Further Research

Although we focus on the risk-reduction property of business longevity, we realize that consumers may use longevity information in various ways. In addition, because of their inherent variability, our domain of interest is services, but goods-oriented firms, particularly food and beverage manufacturers, also use the year of their inception in their advertising. We posit that the primary objective for these firms is to communicate information about the production process, such as the "old-world" technology that some food manufacturers use. In some product categories, older technology may be viewed as more labor intensive, time intensive, or detailed and careful and therefore may be appealing to consumers. Additional research should examine the specific inferences consumers make about product manufacturers that communicate how long they have been in business.

It also would be worthwhile to examine how consumers perceive information about business longevity in rapidly changing markets (e.g., high-tech industries). If technical sophistication represents the key indicator of quality, business longevity may be a liability. In addition, because rapidly evolving industries pose a specific kind of risk to consumers, further work might examine whether such risk affects the impact of business longevity. Thus, additional research should attempt to determine the conditions in which newer firms are perceived as superior.

Finally, our study could be extended to investigate how age influences quality perceptions. We investigate the impact of longevity on an aggregate measure of quality, but specific quality dimensions might be influenced differently (Parasuraman, Zeithaml, and Berry 1985). Therefore, it would be helpful to determine which of the five dimensions of service quality business longevity influences most.

REFERENCES

Alba, Joseph W., Susan M. Broniarczyk, Terence A. Shimp, and Joel E. Urbany (1994), "The Influence of Prior Beliefs, Frequency Cues, and Magnitude Cues on Consumers' Perceptions of Comparative Price Data," *Journal of Consumer Research*, 21 (September), 219–35.

——— and Howard Marmorstein (1987), "The Effects of Frequency Knowledge on Consumer Decision Making," *Journal of Consumer Research*, 14 (June), 14–25.

Anderson K. and R. Zemke (1998), *Delivering Knock Your Socks Off Service*. New York: American Management Association.

Anderson, Norman H. (1965), "Averaging Vs. Adding as a Stimulus-Combination Rule in Impression Formation," *Journal of Experimental Psychology*, 70 (4), 394–400.

Bawa, V., S. Brown, and R. Klein (1979), *Estimation Risk and Optimal Portfolio Choice*. Amsterdam: North-Holland.

Berry, L. and A. Parasuraman (1991), *Marketing Services: Competing Through Time*. New York: The Free Press.

Bettman, James (1973), "Perceived Risk and Its Components: A Model and Empirical Test," *Journal of Marketing Research*, 10 (May), 184–90.

——— (1975), "Information Integration in Consumer Risk Perception: A Comparison of Two Models of Component Conceptualization," *Journal of Applied Psychology*, 60 (June), 381–85.

Blumberg, D. (1991), *Managing Service as a Strategic Profit Center*. New York: McGraw-Hill.

Boulding, W., A. Kalra, and R. Staelin (1999), "The Quality Double Whammy," *Marketing Science*, 18 (4), 463–84.

———, ———, ———, and Valarie Zeithaml (1993), "A Dynamic Process Model of Service Quality: From Expectations to Behavioral Intentions," *Journal of Marketing Research*, 30 (February), 7–27

Brown, S. (1979), "The Effect of Estimation Risk on Capital Market Equilibrium," *Journal of Financial and Quantitative Analysis*, 14 (2), 321–34.

Camerer, Colin and Martin Weber (1992), "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity," *Journal of Risk and Uncertainty*, 5 (4), 325–70.

Campbell, Margaret C. and R C. Goodstein (2001), "The Moderating Effect of Perceived Risk on Consumers' Evaluations of Product Incongruity: Preference for the Norm," *Journal of Consumer Research*, 28 (December), 439–49.

Caves, R. (1998), "Industrial Organization and New Findings on the Turnover and Mobility of Firms," *Journal of Economic Literature*, 36 (4), 1947–82.

Chao, Paul (1989), "The Impact of Country Affiliation on the Credibility of Product Attribute Claims," *Journal of Advertising Research*, 29 (April–May), 35–41.

Coles, J. and U. Lowenstein (1988), "Equilibrium Pricing and Portfolio Composition in the Presence of Uncertain Parameters," *Journal of Financial Economics*, 22 (2), 279–303.

Cox, Donald and Stuart Rich (1964), "Perceived Risk and Consumer Decision Making: The Case of Telephone Shopping," *Journal of Marketing Research*, 1 (November), 32–39.

Crosby, Lawrence A., Kenneth R. Evans, and Deborah Cowles (1990), "Relationship Quality in Services Selling: An Interpersonal Influence Perspective," *Journal of Marketing*, 54 (July), 68–81.

Crum, R., D.J. Laughhunn, and J.W. Payne (1981), "Risk-Seeking Behavior and Its Implications for Financial Models," *Financial Management*, 5 (Winter), 20–27.

Dodds, William B., Kent B. Monroe, and Dhruv Grewal (1991), "Effects of Price, Brand, and Store Information on Buyers' Product Evaluations," *Journal of Marketing Research*, 28 (August), 307–319.

Dowling, G.R. and Richard Staelin (1994), "A Model of Perceived Risk and Intended Risk-Handling Activity," *Journal of Consumer Research*, 21 (June), 119–34.

Erdem, Tulin (1998), "An Empirical Analysis of Umbrella Branding," *Journal of Marketing Research*, 35 (August), 339–51.

Fishburn, P.C. (1977), "Mean-Risk Analysis with Risk Associated with Below-Target Returns," *American Economic Review*, 67 (2), 116–26.

Kahneman, D. and Amos Tversky (1979), "Prospect Theory: An Analysis of Decisions Under Risk," *Econometrica*, 47 (2), 262–91.

Kirmani, Amna (1990), "The Effect of Perceived Advertising Costs on Brand Perceptions," *Journal of Consumer Research*, 17 (September), 160–71.

Klein, R.W. and V.S. Bawa (1977), "The Effect of Estimation Risk on Optimal Portfolio Choice," in *Estimation Risk and Portfolio Choice*. New York: North-Holland, 50–63.

Laughhunn, D J., J.W. Payne, and R. Crum (1980), "Managerial Risk Preferences for Below-Target Returns," *Management Science*, 26 (12), 1238–49.

Lynch, John G., H Marmorstein, and M.F. Weigold (1988), "Choice from Sets Including Remembered Brands: Use of Recalled Attributes and Prior Overall Evaluations," *Journal of Consumer Research*, 15 (September), 169–84.

Mittal, Vikas, William T. Ross, and Michael Tsiros (2002), "The Role of Issue Valence and Issue Capability in Determining Effort Investment," *Journal of Marketing Research*, 69 (November), 455–68.

Miyazaki, Anthony, Dhruv Grewal, and R.C. Goodstein (2005), "The Effect of Multiple Extrinsic Cues on Quality Perceptions: A Matter of Consistency," *Journal of Consumer Research*, 32 (June), 146–53.

Monroe, K. and W. Dodds (1988), "A Research Program for Establishing the Validity of the Price-Quality Relationship," *Journal of the Academy of Marketing Science*, 16 (1), 151–68.

Mowen, John (1987), *Consumer Behavior*. New York: Macmillan.

Murray, Keith B. (1991), "A Test of Services Marketing Theory: Consumer Information Acquisition Activities," *Journal of Marketing*, 55 (January), 10–25.

Parasuraman, A., Valarie A. Ziethaml, and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49 (Autumn), 41–50.

Payne, John, D. Laughhunn, and R. Crum (1980), "Translation of Gambles and Aspiration Level Effects in Risky Choice Behavior," *Management Science*, 26 (10), 1039–1060.

———, ———, and ——— (1981), "Further Tests of Aspiration Level Effects in Risky Choice," *Management Science*, 27 (8), 953–58.

Petty, R. and J. Cacioppo (1986), *Communication and Persuasion: The Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Purohit, D. and J. Srivastava (2001), "Effect of Manufacturer Reputation, Retailer Reputation, and Product Warranty on Consumer Judgments of Product Quality: A Cue Diagnosticity Framework," *Journal of Consumer Psychology*, 10 (3), 123–34.

Richardson, Paul S., Alan S. Dick, and Arun K. Jain (1994), "Extrinsic and Intrinsic Cue Effects on Perceptions of Store Brand Quality," *Journal of Marketing*, 58 (October), 28–36.

Roberts, John and Glen Urban (1988), "Modeling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice," *Management Science*, 34 (2), 167–85.

Rust, R., J. Inman, J. Jia, and A. Zahorik (1999), "What You Don't Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions," *Marketing Science*, 18 (1), 77–92.

Sengupta, J., R.C. Goodstein, and D. Boninger (1997), "All Cues Are Not Created Equal: Obtaining Attitude Persistence Under Low-Involvement Conditions," *Journal of Consumer Research*, 23 (March), 351–61.

Schiffman, Leon and Leslie Kanuk (1994), *Consumer Behavior*. Englewood Cliffs, NJ: Prentice Hall.

Shimp, Terence A. and Willam O. Bearden (1982), "Warranty and Other Extrinsic Cue Effects on Consumer Risk Perceptions," *Journal of Consumer Research*, 9 (June), 38–46.

Simon, Herbert (1955), "A Behavioral Model of Rational Choice," *Quarterly Journal of Economics*, 69 (1), 99–118.

Slovic P. and S. Lichtenstein (1971), "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment," *Organizational Behavior and Human Performance*, 6 (6), 649–744.

Stambaugh, R. (1997), "Analyzing Investments Whose Histories Differ in Length," *Journal of Financial Economics*, 45 (3), 285–331.

Van Trijp, Hans, Wayne Hoyer, and J. Jeffrey Inman (1996), "Why Switch? Product Category–Level Explanations for True Variety-Seeking Behavior," *Journal of Marketing Research*, 33 (August), 281–92.

Zeithaml, V., A. Parasuraman, and L. Berry (1990), *Delivering Quality Service: Balancing Customer Perceptions and Expectations*. New York: The Free Press.

Alan G. Sawyer, Juliano Laran, & Jun Xu

The Readability of Marketing Journals: Are Award-Winning Articles Better Written?

This is a study of the readability of articles in four marketing journals: *Journal of Marketing, Journal of Marketing Research, Journal of International Marketing,* and *Journal of Public Policy & Marketing.* For each journal, the authors compare articles that have won an award with articles that have not. The authors find that award-winning articles are more readable, as measured by indexes focusing on sentence and word length, than nonwinning articles. The authors also identify and analyze other characteristics of more readable journal articles and discuss the importance of good writing.

Keywords: readability, award-winning articles, marketing journals, path analysis

The October 2005 *Journal of Marketing* (*JM*) includes 11 essays about challenges to marketing and the content and impact of published research in marketing journals. The essay "Readability and the Impact of Marketing" (Bauerly, Johnson, and Singh 2005) documents a steep decline of readability in *JM* articles in the early 1970s, when marketing academics became more scientific and specialized. Since then, readability has remained low. This article offers an expanded empirical investigation of the readability of marketing journal articles.

The current manuscript evaluation criteria for *JM* make it clear that readability is important. "All manuscripts are judged not only on the depth and scope of the ideas presented and their contributions to the field, but also on whether they can be read and understood. Keep sentences short so the reader does not get lost before the end of a sentence.... The journal is designed to be read, not deciphered" (American Marketing Association [AMA] 2005).

Despite frequent calls for improved readability, many marketing journal editors view writing in academic journals as badly in need of improvement. As Winer (1998, p. iv) states in a *Journal of Marketing Research* (*JMR*) editorial, "Authors consistently underestimate the importance of writing to a paper's ultimate disposition. Reviewers are justifiably impatient with manuscripts in which the contribution is not clear, there are many typos (even in this era of spelling checkers), and so on. Do not believe that these kinds of problems will be eliminated through revisions; if the writing is sufficiently unclear, the author might not get that opportunity." Finally, in his outgoing *Journal of Consumer Research* editorial, Lutz (1990, p. 244) states, "Frankly, it is appalling to see how badly so many of us write." In addition to bad writing, virtually all editors focus on the excessive length of submitted manuscripts. "In the vast majority of cases, manuscripts are simply too long relative to their potential contribution.... Ceteris paribus, the shorter the article, the more likely it is that someone will actually read it" (Lutz 1990, p. 245).

Clear scientific writing is essential for research to be understood, published, and impactful. As Mick (2005, p. 1) states, "The unavoidable reality is that writing effectively is not just a necessary condition for getting published. Better writing propels an academic's influence and reputation, and the finest writing is more often found among the most distinguished researchers." Moreover, good writing is essential "if marketers want to communicate across specializations, across functional areas, and outside the marketing discipline and if marketing science is to influence practice" (Bauerly, Johnson, and Singh 2005, p. 19).

Research Objectives

The prime research objective was to test the validity of Mick's (2005) logical speculation that impactful articles written by the most distinguished researchers are more readable. Readability is defined as "the ease of understanding or comprehension due to the style of writing" (Klare 1963, p. 15) and "the extent to which (a group of) readers understand it, read it at an optimal speed, and find it interesting" (Dale and Chall 1949, p. 19). We used articles that

---

Alan G. Sawyer is J.C. Penney Professor of Marketing (e-mail: alan.sawyer@cba.ufl.edu), and Juliano Laran (e-mail: laran@ufl.edu) and Jun Xu (e-mail: jun.xu@cba.ufl.edu) are doctoral candidates in Marketing, University of Florida. The senior author thanks the Warrington College of Business and the J.C. Penney Foundation for financial assistance with this research. Roger Calantone of Michigan State University and Dwayne Ball and Daniel Chen of the University of Kansas provided valuable assistance with the principal components analysis and path analysis, for which the authors are indebted. Taking their own advice, the authors asked Lita Kurth to copyedit one of their "final" drafts, and they are grateful for her improvements to the manuscript. They also appreciate the helpful advice of the anonymous *JM* reviewers. All three authors contributed equally to this work.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

had won a major award in an AMA academic journal as the benchmark of distinguished research. Another research objective was to identify article and author characteristics that may contribute to or at least correlate with readability. Bauerly, Johnson, and Singh (2005) speculate that a reason for the decline in readability is that more articles are written by academics from more methodologically rigorous and theoretically based doctoral programs. This speculation seems valid, and several other aspects of the authors' backgrounds and the content of the articles may be related to better readability. Because Bauerly, Johnson, and Singh did not code the contents of analyzed articles, they were unable to examine correlates of readability other than the publication date.

In our research, we code the content of sampled articles and test the relationship of readability to various article types and author characteristics. We use these characteristics to control statistically for differences among award-winning and nonwinning articles. The content of the article and other such characteristics might influence readability. We predict better readability for some types of empirical articles than nonempirical ones for three reasons. First, empirical articles, particularly experiments, have a prototypical report format that might enable briefer, more readable sentences. Second, nonempirical articles focusing on managerial, theoretical, or other complex issues might contain longer and more complex sentences with longer words. A particular type of nonempirical article, the literature review, might be particularly less readable. Third, empirical articles with data and statistics might be rated as more readable because most contain more numbers. This could result in an artifactual relationship because numbers are coded as one-syllable words with only one or a few characters. We had no expectation regarding whether different journals would be rated as more readable, other than the variation in content or author experience. Articles in more prestigious academic journals might be more readable if, for example, authors write more carefully and seek more advice for manuscripts targeted at these journals. However, articles in journals with a more practitioner-oriented audience may be more readable if, in consideration of those readers, authors write with simpler sentences and less jargon.

Author characteristics may be important. For example, more experienced authors likely write better, regardless of whether the article has won an award. Top researchers likely think more clearly, which is an essential skill for good writing. Award-winning authors are likely to be more experienced and have an advantage of more practice and more feedback from editors and colleagues on more manuscripts. Coauthorship is also possibly related to readability. Coauthored articles might benefit in terms of readability because each coauthor would try to improve the writing of the other authors or would work on assigned sections about which they are more knowledgeable and thus write better. Conversely, coauthoring may lead to coordination problems that could result in less clear, more cumbersome writing. Hartley, Pennebaker, and Fox's (2003) literature review shows mixed results for the readability of coauthored articles, and their own empirical investigation finds no differences between single-authored and coauthored works.

Characteristics of Good Writing

Scientific writing presents several challenges. Complex topics may require complex writing to convey accurate information, with frequent use of jargon, qualifying phrases, data, statistics, footnotes, and references. In addition to demanding a correct presentation of all the facts, readers "won't put up with being delayed, lost, confused, bored or taken down unnecessary trails by failures of craft" (Zinsser 1988, pp. 57–58). The many dimensions of good scientific writing include having good grammar and sentence structure; using appropriate, short, and easy-to-understand words and sentences; avoiding unnecessary technical jargon; and keeping the length proportionate to the likely contribution. Other imperatives beyond the basics are probably more important and subtle (Sawyer 1988). These include properly positioning and "selling" the contribution of the research (Peter and Olson 1983), writing for a specific audience and journal, and using an interesting writing style with a clear and appropriate structure (Holbrook 1986).

Although existing objective methods do not measure the more subtle aspects of good writing, several formulas that calculate readability indexes are available to assess the basics (for a summary of this research, see DuBay 2004). These formulas include the length of words, as expressed by the number of syllables or characters, and the length of sentences, as expressed by the number of words or long words. These basics do not guarantee a readable manuscript; the formulas "measure only 'surface features' of text and ignore content and organization" (DuBay 2004, p. 35). However, although sentence and word length do not *cause* readability difficulty, they are highly *correlated* with this problem and thus can be used to suggest possible improvements.

Attention to readability indexes can help authors improve their writing. As Judith Swan (qtd. in Knight 2003, p. 377) stated, "Jargon is less pernicious if you understand what is going on." The structure of the sentence is more apt than long words to be the prime readability problem (Gopen 2004; Gopen and Swan 1990). Readers persist longer with easier-to-read material and find it more interesting if the writing matches their preferred reading level.

Readability indexes are not without problems (Redish and Selzer 1985). For example, shorter sentences and words are not always clearer than longer ones, and readability does not measure the appropriateness or accuracy of the content. Moreover, although scientific texts with many scientific terms may have lower readability ratings, they may be easier to read for those trained in subspecialties, who may even prefer jargon or are bored with simpler, "more readable" text.

Despite these issues, readability formulas have received corroboration from initially skeptical cognitive psychology researchers, who have tried to go beyond simple measures of word and sentence arithmetic. Cognitive theorists who study how readers learn posit that readers construct meaning because meaning is not in the words by themselves on the page. Readers store information in long-term memory, link new information to prior knowledge, and use metacogniton to think about and plan the learning process (DuBay 2004; Gopen and Swan 1990). After examining the empirical effects on comprehension of cognitive and structural text elements, such as coherence and organization, several cognitive theorists have reversed their negative opinions about readability formulas (see DuBay 2004, p. 35). For example, Kintsch and Miller (1984, p. 222) conclude that "these formulas are correlated with the conceptual properties of text" and are the strongest predictors of reading speed and comprehension difficulties.

Readability of Top Journals and Articles

Bauerly, Johnson, and Singh (2006) compared the abstracts of 70 *JM* articles with 200–400-word sections of their introductions. The abstracts were rated significantly more readable than the material from the articles, though the two scores were significantly correlated. Bauerly, Johnson, and Singh concluded that the authors knew how to write better than they did in the text. They speculated that a reason for the poorer writing in the text is that some researchers might write less simply than desirable and overuse technical jargon to impress readers. Armstrong (1980) conducted two tests of this possible phenomenon. In the first test, he found that the readability was inversely related to ratings of journal prestige. In a second, better-controlled test, he took four excerpts from marketing journals and manipulated their reading difficulty by changing the length of sentences, varying the difficulty of words, and eliminating/adding unnecessary words. Thirty-two faculty members judged the less readable excerpts with the same content to be more competent. Hartley, Trueman, and Meadows (1988) found the same result with psychology articles. Armstrong (1989) subsequently reanalyzed both data sets and concluded that more easily read articles were rated less prestigious ("If I can read this easily, it must not be prestigious"), but beyond some point, differences in readability had no further effect.

Additional research has found counterintuitive relationships between readability and factors such as the number of citations of the article and rejection of a manuscript by a journal. Stremersch, Verniers, and Verhoef (2007) analyzed four top-rated marketing journals and found a small but statistically significant, negative relationship between readability and the number of times an article had been cited. They concluded (p. 1) that "the readability of an article may hurt academic credibility and impact." Articles with fewer citations were more readable. A similar negative relationship to readability was reported by Metoyer-Duran (1993), who analyzed the readability scores of articles accepted and rejected for publication in a library science journal. Readability scores of rejected manuscripts were significantly better than those of accepted ones. In contrast to the negative results of Armstrong (1980) and others, Oppenheimer's (2005) experimental research showed greater preference for more readable text. He conceptualized that because of greater processing fluency, texts with simpler words should result in more positive judgments. The results of five highly controlled experiments were consistent with this prediction and his hypothesized explanation. Respondents judged the authors of texts with simpler words to be more intelligent, and this effect was mediated by ratings of processing fluency. Direct manipulations to increase the fluency of texts resulted in judgments that were more positive. Only when the source of fluency was obvious did respondents discount their reliance on the fluency cue.

Additional correlational evidence from research in psychology suggests that articles that are considered more important are more readable than their less important counterparts on the standard readability indexes. Hartley, Sotto, and Pennebaker (2002) reported four correlational studies that took an excerpt from the introductory pages of each article and scanned it into a word-processing package. The first study compared the readability scores for excerpts from 19 of the most influential journal articles in cognitive science with 19 articles that were not cited as influential but were published in the same journal issues. The impactful articles had significantly better readability scores, and there was a positive .41 rank correlation between the Flesch scores and the position of 35 articles within the 100 most influential publications (which also included books and monographs). Hartley, Sotto, and Pennebaker's second study found that 36 articles from *Forty Articles That Changed Psychology* (Hock 1992) had significantly better readability scores and shorter sentences than 36 counterpart articles not cited by Hock. However, two other studies by Hartley and colleagues found no relationship between readability and the number of citations.

In summary, the evidence is mixed regarding whether articles written in a less readable way are preferred, more likely to be accepted at a journal, associated with greater impact, considered more prestigious, or written by more intelligent authors. Most relevant for the purposes of this study is the evidence that renowned articles may be more readable.

Research Method

This research assessed whether award-winning marketing articles in four academic journals published by the AMA are more readable than counterpart articles that have won no awards. We also examined variables that may characterize more readable publications. To avoid problems with the validity of sampling only small, single excerpts from articles and avoid errors between scanned and original texts, we copied almost all of each article's text downloaded from Internet journal databases available in HTML format into Microsoft Word. Unlike prior research, our study used multiple measures of readability.

Sample

Journals. The sample included award-winning articles in *JM*, *JMR*, *Journal of International Marketing* (*JIM*), and *Journal of Public Policy & Marketing* (*JPP&M*). There were four reasons for choosing this particular set of journals. First, these academic journals are the only ones published by the AMA. Second, online versions of their articles published since 1990 are available in HTML format. Third, these journals differ considerably and thus provide within-study replication across a diverse set of journals. In particular, *JPP&M* and *JIM* vary greatly from *JM* and *JMR* in terms of their focus, impact, and prestige. The *Web of Sci-*

ence Journal Citation Reports impact index (based on the number of citations of an average article in a journal within two years of publication) and Baumgartner and Pieters's (2003) structural influence index (based on the proportion of citations in other journals of all articles published in that journal) show that the recent impact and overall influence were high for both *JM* (4.1 and 19.1, respectively) and *JMR* (2.6 and 16.4) and lower for both *JPP&M* (1.2 and .8) and *JIM* (.4 and .2). Moreover, academics vary widely in their rankings of these journals. Two studies (Bauerly and Johnson 2005; Hult, Neese, and Bashaw 1997) show large differences in the ranking of *JM* (1st in both studies) and *JMR* (3rd and 2nd, respectively) compared with *JPP&M* (36th of 54 journals and 12th of 34 marketing journals) and *JIM* (unranked and 26th of 34).

The fourth, and most important, reason for choosing these four journals is that each presents annual awards for outstanding articles. *Journal of Marketing* gives two awards for articles published during the previous year: the Harold H. Maynard Award for the best article on "marketing theory and thought" and the Marketing Science Institute/H. Paul Root Award for the "best contribution to marketing practice."[1] *Journal of Marketing Research* also gives two awards: the William F. O'Dell Award for the article published five years previously that made "the most significant, long-term contribution to marketing theory, methodology, and/or practice" and the Paul E. Green Award for the article published during the previous year that "shows or demonstrates the most potential to contribute significantly to the practice of marketing research and research in marketing." *Journal of International Marketing* also gives two awards: the S. Tamer Cavusgil Award for an article from the previous year that made "the most significant contribution to the advancement of the *practice of international marketing management*" and the Hans B. Thorrelli Award for the article published within the past five years that has made the "most significant, long-term contribution to *international marketing theory or practice.*" *Journal of Public Policy & Marketing* gives the Thomas C. Kinnear Award for "significant contribution to the understanding of marketing and public policy issues" within the past three years.

Nonwinning articles. For each award-winning article, we chose a counterpart nonwinning article from the same journal issue to control for temporal issues, such as editors and reviewers. We arbitrarily set the nonwinner as the penultimate article in the issue; we considered the last article the one immediately preceding any notes or rejoinders. The reason for choosing nonwinning articles positioned near the end of the issue was to attempt a stronger "manipulation" of article quality. Many editors place articles they view as more important toward the front of the issue (e.g., Medoff 2003). If the beginning article was an editorial, we coded the article immediately following any editorials as the first. In some years, there were multiple recipients of a given award or articles that received multiple awards. We counted articles that won multiple awards (e.g., both the Green and the O'Dell awards in *JMR* in a given year) only once with only one counterpart nonwinner. If different award-winning articles appeared in the same issue, the article appearing just before the penultimate one became the second nonwinner. For example, if there were 11 articles in an issue containing an award winner and the final 3 were research notes, the nonwinner would have been the seventh article. If the same issue contained a second award-winning article, its nonwinning counterpart would have been the sixth. The final sample of 81 award-winning articles included 32 from *JM*, 22 from *JMR*, 15 from *JIM*, and 12 from *JPP&M*. These articles included 13 Maynard Award winners, 15 MSI/Root Award winners, and 4 that had won both awards in *JM*; 11 O'Dell Award winners, 9 Green Award winners, and 2 that had won both awards in *JMR*; 7 Thorrelli Award winners and 8 Cavusgil Award winners in *JIM*; and 12 Kinnear Award winners in *JPP&M*. If we include the 81 nonwinners from the same issues as the winners, the sample totaled 162 articles.

Measures

After we copied each entire article into Microsoft Word, we trimmed the text of each article of the title and abstract; all headings and subheadings; any mathematical equations (note that some HTML versions we used did not include the full text of any equations); quotations that included more than one sentence; and all tables, figures, and footnotes. We uploaded the remaining text to a Web site (Taylor and Intuitive Systems 2006) that calculated several indexes of readability that include word length defined in various ways (e.g., characters or syllables per word, words longer than a certain number of syllables or characters) and the length of sentences (number of words or words of a particular length per sentence). Four readability scores, normed to the school grade appropriate for that reading difficulty, were the Flesch–Kincaid (hereinafter, the Kincaid score), the automated readability index (ARI), the Gunning–Fog index (Fog), and the SMOG grade-level scores. We also calculated the Flesch index and the Lix index, which vary between 1 and 100. (The Appendix contains the formulas for all indexes.) Other readability measures produced by the Taylor and Intuitive Systems (2006) Web site included the number of characters in the manuscript, the number of words, their average length in terms of the number of characters and syllables, the number of sentences, and their average length.

In addition to the individual indexes, we calculated an average score, which we labeled as "average readability grade," from the Kincaid, ARI, Fog, and SMOG scores (coefficient $\alpha = .98$). Our readability results focus on the average readability grade level and three individual readability scores: the Flesch index (used by Bauerly, Johnson, and Singh 2005; Hartley, Pennebaker, and Fox 2003), the Lix score, and, though it is included in the average readability grade score, the Kincaid grade score presented by itself, because some readability researchers consider it the most appropriate one for technical documents (DuBay 2004). We also report other, perhaps more interpretable measures, including the length and number of sentences and words.

---

[1] A third *JM* award was established in 2001 by the Sheth Foundation for the *JM* article published 6–10 years prior with the largest contribution to marketing. Our analysis did not include Sheth Foundation Award articles, because all but one also won one of the two other *JM* awards.

Other variables in the data set were potential correlates of readability. These included the number of authors, the award, the publication year, the issue number, the position of the article in the issue, the number of references within the article, the number of citations of the article, and its status as either an award winner or nonwinner. The number of references within the article may serve as a surrogate for a literature review article. To control for the length of the manuscript, which might be correlated with the number of references used, we divided the number of references within the article by the number of sentences. We obtained information about publications and citations from the Web of Science Social Science database. We coded the number of publications in six marketing journals (the four AMA journals included in our research plus *Journal of Consumer Research* and *Marketing Science*) for the most published author and the combined total publications of all authors. We divided the number of citations of each article by the number of years in print to create an annual citations variable.

Finally, we coded three other article characteristics. The first was whether the article included equations before being trimmed for analysis. Of the articles, 36% contained equations in the original text. A second variable, data source, included subcategories of data from secondary sources (30%), from experiments (14%), or from authors' surveys (41%). In addition, 26% were nonempirical and thus contained no data.[2] A third variable category included codes for whether the article topic was methodological (25%), managerial (57%), behavioral (17%), or theoretical (14%). We coded these subcategories as dummy variables. Some categories within both the data source and topic variables were highly negatively correlated. Even after we dropped one category so that any analysis would not be overidentified, some subcodes, such as methodological and managerial, remained nearly mutually exclusive. Thus, we kept only the data source and topic classifications with small, nonsignificant intercorrelations. The final set of five article characteristics included whether the article (1) contained equations, analyzed data from (2) experiments or (3) secondary sources, and focused on (4) methodological or (5) behavioral issues.

We then factor-analyzed these five article characteristics with the recommended M-plus algorithm (Muthén and Muthén 2004), which calculates tetrachoric correlations for a confirmatory factor analysis (CFA) for dichotomous variables instead of Pearson phi correlations. The results were two orthogonal principal components. The first component, which we labeled as "behavioral experiments," had a .91 correlation with the behavioral and experimental variables, and the second component, which we labeled as "methodological secondary data models," had a .83 correlation with each of the other three article characteristics. The fit of the CFA model was excellent ($\chi^2 = 8.32$, d.f. = 5; comparative fit index = .98; and root mean square error of approximation = .064). We used these two components in further analyses that included descriptions of article characteristics.

Results

Overall Readability

The average readability grade score across all articles averaged 16.2, which is considered difficult reading for people below the grade of a college senior. In comparison, the average readability grade levels for the *New York Times* and *PC World* are 7.6 and 11.7, respectively (Taylor and Intuitive Systems 2006). The Flesch readability score, which has been widely used in prior research and is the only index for which higher scores indicate better readability, averaged 35.3 for these 162 journal articles, compared with average scores of 70.4 and 59.4, respectively, for the *New York Times* and *PC World* (Taylor and Intuitive Systems 2006). Of the 15 articles with the best readability grade scores, 13 were winners, with scores ranging from 12.3, suitable for a high school senior, to 14.4, suitable for a sophomore in college. Of the 11 articles rated least readable, 9 were nonwinners, with the readability grade scores ranging from presumably graduate school levels of 21.3 to 18.3. Table 1 shows the scores for the 5 articles rated most and least readable.

Main Results

The prime goal of this research was to test whether award-winning articles in marketing journals were more readable than nonwinners. We used readability grade, as well as other variables, to define readability. We also analyzed the relationships between winning articles and the length of words and sentences. We defined word length as the number of characters and sentence length as the number of words. The second research objective was to search for any article characteristics in addition to word and sentence length that distinguish more readable articles. Specifically, we examined three article characteristics: the two principal component scores from the CFA (behavioral experiments and methodological secondary data models) and the number of references per sentence for each article.

To test these two objectives and assess various relationships simultaneously, we used a path analysis that examined the relationships among single measures of each of seven variables: articles with behavioral experiments and articles focusing on methodological secondary data models, the number of references per sentence, the length of both words and sentences, readability grade, and whether the article had won an award. The results of the path analysis, which we performed with LISREL, showed an excellent fit ($\chi^2 = 12.35$, d.f. = 10, $p < .25$; comparative fit index = .99; and root mean square error of approximation = .039). The respective standardized beta coefficients and z-values appear in Figure 1.

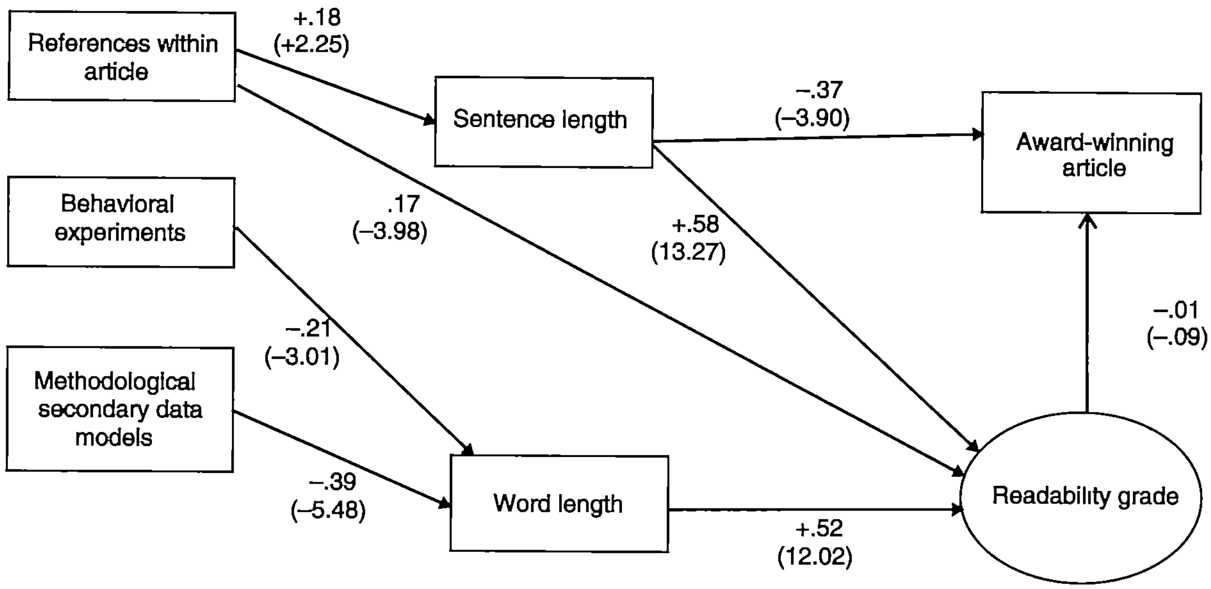
---

[2] Because some articles were coded into more than one category, these percentages add to more than 100%. For example, one article, which analyzed data from field experiments previously conducted over many years by a commercial research firm, was coded in both the experiment and secondary data categories.

TABLE 1
Articles Rated Most and Least Readable by Average Readability Grades

Authors	Award Winner	Readability Grade	Flesch Score	Sentence Words
Most Readable				
Golder and Tellis (1993)	Yes	12.3	50.7	18.2
Desiraju and Shugan (1999)	Yes	12.5	53.0	20.3
Stremersch and Tellis (2002)	Yes	13.2	47.5	19.8
Robinson and Min (2002)	No	13.6	44.0	19.5
Hoch, Dreze, and Purk (1994)	Yes	13.7	48.0	22.3
Least Readable				
Vorhies and Morgan (2003)	No	21.3	11.4	31.1
Prasad, Ramamurthy, and Naidu (2001)	No	20.2	18.5	31.5
Morgan, Kaleka, and Katsikeas (2004)	No	19.7	19.4	30.1
Glazer (1991)	Yes	19.6	24.9	28.4
Gencturk and Kotabe (2001)	No	19.3	20.8	29.5

Figure 1
Causal Path of Article Characteristics on Readability Grade



Notes: All coefficients are statistically significant at $p <$ 05, except for the null relationship between "Readability" and "Award-winning article" Because lower grades indicate better readability, a negative coefficient indicates a positive relationship to readability. For example, longer words and sentences cause higher readability grades (i.e., poorer readability) The top value beside each path is the size of the standardized relationship. The resulting Z-test is in parentheses.

Award winners versus nonwinners. An analysis of covariance (ANCOVA) with the three leftmost covariates in Figure 1 and journal and winner as fixed factors augmented the path analysis. The ANCOVAs indicated that every readability index in the Appendix was significantly better for award-wining articles ($p < .02$). Word length was related to readability but not to award-winning status. Sentence length was related to both readability and whether the article won an award.

However, when we controlled for all of the variables in the model simultaneously in the path model, there was no direct relationship between readability grade and whether an article won an award. As Figure 1 shows, the path analysis indicated that sentence length was the only component of readability grade that distinguished award winners from nonwinners. This was presumably due to the fact that the relationship between sentence length and readability grade was larger (.58) than its relationship to whether the article had won an award (–.37), thus nullifying the relationship. A second path analysis that eliminated the path from sentence length to winner showed a significant path (b = –.25, p = .003) of the same size of the bivariate correlation between readability grade and winning articles. However, further analysis that compared these two models revealed that the model in Figure 1 with the direct (but nonsignificant) path from readability to winner fits the data better than the one without this path ($\Delta\chi^2 = 12.64$, Δd.f. = 1, $p < .05$). It is noteworthy that the model in Figure 1 replicated across other measures of readability. Corresponding path analyses with the Flesch, Lix, or Kincaid scores as the definition of readability produced results with statistically significant paths of essentially the same size as when we used the average readability grade.

A possible counterexplanation for the observed differences between award-winning and nonwinning articles might be that numbers are counted as words of only a few characters and syllables. However, this cannot explain the greater readability of award winners for several reasons. First, word length did not differ between winners and nonwinners; only the length of sentences differed. Second, we included the article characteristics to control for such possible confounds. Third, all the *JMR* articles, winners and nonwinners, were empirical, but there were significant differences in readability. Thus, empirical status and the possibility of an artifactual result due to the way the readability software coded the word length of numbers cannot explain the difference between winners and nonwinners.

Readability scores for the four journals. A reason for the ANCOVAs was to test for differences among the journals in which the article was published. The results showed no meaningful differences among the four journals or presumed differences in prestige or impact. That there were no significant interactions between award winners and journal suggests that the results between award winners and nonwinners replicated across journals.

Characteristics of more readable articles. All the relationships specified in the paths in Figure 1 were statistically significant, except for the nil one between readability grade and whether the article had won an award. The two principal component scores used to summarize article types were significantly related to word length. As Figure 1 shows, there were shorter words in articles that reported experiments with behavioral conceptualization (b = –.21) and those that focused on methodological models analyzing secondary data (b = –.39). The only direct relationship of an article characteristic to sentence length was the number of references per sentence (b = +.18), which was also associated with worse readability grade levels (b = +.17). Thus, sentence length only partially mediated the effects of the number of references on readability. Longer sentences resulted in worse readability grades (b = +.58), as did longer words (+.52). These latter two results are near tautological because each of these two variables, in some form, is included in all the readability formulas.

Other results. Additional analyses showed that, consistent with Mick's (2005) logical conjecture, articles written by authors who have published more articles had better readability scores ($r = -.27$, $p < .001$). These experienced authors were also much more likely to write award-winning articles ($r_{\text{number of publications, award winner}} = .38$, $p < .001$). However, consistent with Hartley, Sotto, and Pennebaker's (2002) results and counter to the negative relationship that Stremersch, Verniers, and Verhoef (2007) find, the total number of an article's citations, the number of an article's citations per year, and coauthorship were not related to readability scores.

Discussion

Award-winning articles in AMA academic journals were more readable than nonwinners on several measures of readability. Readability, which includes measures of the length of words and sentences, was not directly influenced by any article characteristic, with one exception: the number of references. The number of references per sentence was associated with longer sentences, but sentence length did not completely mediate the effects of the number of references on readability. Thus, theoretical position and literature review articles, which often contain more references, were more likely to be rated as less readable beyond the merely tautological need for longer sentences to include the increased percentage of references. There were shorter words in articles that focused on behavioral experiments and methodological mathematical models that used secondary data than in other articles, such as articles on managerial topics, nonempirical articles, and articles that used survey data. However, word length completely mediated the effects of these various types of articles on readability.

Overall, these results show that with the possible exception of review articles, which may be less readable because of longer sentences, no particular type of article was more or less readable. The key is the author's discipline in writing clearly; the use of shorter words and shorter sentences can also help. According to Ross-Larson (1999), any sentence with more than 25 words is likely too long. Sentences of the articles in our sample averaged 25.1 words. In the classic *Elements of Style*, Strunk and White (2000, p. 23) succinctly advise to "omit needless words." Our analyses of the readability scores of journal articles support this advice. Sentence length was the key factor. Good sources for details about how to shorten sentences include Cook (1985), Flesch and Lass (1996), and Ross-Larson (1996, 1999). In addition to eliminating needless words, phrases, and sentences, Ross-Larson (1999, p. 18) suggests that writers "read a sentence aloud. If they stumble, gasp for air, the sentence is not well crafted, and the stumbles and gasps show them where to make repairs."

Brevity is a good thing, but it is not the prime goal of good writing. The key is improved reading ease and comprehension. Psychological research (e.g., Kintsch and Miller 1984) has shown that these factors are most highly correlated with readability, as measured by the formulas in the Appendix. Taken literally, winners and nonwinners differed by an average of less than a school grade. Although this difference may not be significant for academic researchers, DuBay's (2004, p. 49) review of readability research concludes that "even small differences in comprehension pay off." Small improvements in writing can help even expert readers read a manuscript faster, while understanding and remembering more.

Moreover, marketing is an applied discipline. Practitioners are more than half of the subscribers of AMA journals (73% for *JIM*, 68% for *JM*, 62% for *JMR*, and 53% for

TABLE 2
Differences Between Award-Winning and Nonwinning Articles

Measure	Award-Winning Articles	Nonwinners
Average readability grade	15.9	16.5
Kincaid grade level	14.2	14.9
Flesch index[a]	37.0	33.9
Sentence length (words)	24.0	26.5
Word length (characters)	5.3	5.4
Total number of words	9060	7203

[a]The Flesch score is the only readability index for which higher scores indicate better readability. It is not included in the mean readability grade scores.

Notes: All means are adjusted by three covariates in ANCOVA: the number of references per sentence in the article, whether it was more apt to be a behavioral experiment, and whether it was more apt to be a methodological modeling article that contained equations. The difference between winning and nonwinning articles for all but word length was statistically significant at $p < .006$. In their order in Table 2, effect sizes, as estimated by omega-squared, of the significant differences were .058, .063, 050, .158, and .195.

JPP&M). There is wide agreement that academics and practitioners do not communicate well (e.g., Crosier 2004). Several essays in the October 2005 issue of *JM* bemoan the lack of impact of academic research on marketing practice, and the writing style of many academic research articles may account for more than a little of this communication gap. If possible, authors should seek a review by a practitioner member of the target audience.

Although the statistically significant differences in the readability grade scores between award winners and nonwinners may seem small, the effect sizes (detailed in Table 2) approximated or exceeded what Cohen (1988; see also Sawyer and Ball 1981) defines for behavioral research as a medium size (6.25% explained variance). This difference is noteworthy because of the many other reasons readability scores might vary. Our comparison of winning and nonwinning published articles was a tough test of the discriminatory power of the readability indexes. There was a strong possibility of no significant differences among articles within top-level academic journals. We compared award winning articles with nonwinning articles, but the nonwinners were not "losers." They were published in AMA journals, which are selective and have high rejection rates.

Interpretation of the implications and importance of the practical effect size of the components of readability scores might be more enlightening than the scores themselves. Award-winning articles contained 2.5 fewer words per sentence (9.4%) than nonwinning articles. Given the 9900 words per award-winning article in *JM*, this 934-word advantage would amount to more than three manuscript pages and close to an average *JM* journal page—enough to accommodate an additional article in an issue that averages 10–11 articles per issue.[3]

Conclusion

Motivation and perspiration are important ingredients for good writing. Although it is well known that multiple drafts are needed, many authors underestimate what is required for a well-written manuscript. Comments from colleagues and copy editors are imperative. Failing to maximize an article's readability is foolish and risky, leads to wasted time for journal reviewers, and increases the chances of rejection. Even after many drafts, authors should take a "final" draft and, one last time, try to simplify, shorten, and improve the structure of sentences and the manuscript as a whole. As Winer (1998, p. iv) advises, "consider the version you are thinking of sending as the n – 1st version. Give it to colleagues to read not only for content but for clarity as well."

We hope that our focus on good writing and the use of precise, short sentences to improve readability in marketing will motivate readers to work even harder in this regard. Although authors should not naively write simply to maximize readability scores, such scores are a useful screen—a "canary in the mineshaft" (Bauerly, Johnson, and Singh 2006, p. 223)—to predict reading ease.[4] Most manuscripts and sentences can be shortened and improved. As Holbrook (1986, p. 106) colorfully states, "Remember that brevity is the soul of wit. In each section, paragraph, and sentence of your paper, shun excessive length as diligently as you would avoid paying extra interest expenses on your credit card. Say exactly what you need to say, no more, then stop. Like this."

---

[3] The technical editor at the AMA estimated that a typical manuscript page has slightly more than 300 words. Journal pages average approximately 930 words in *JMR*, approximately 1060 words in *JM* and *JPP&M*, and approximately 500 words in *JIM* (personal communication, 2005).

[4] After the elimination of the title, the abstract, the headings, the references, the tables, the figure, the footnotes, the quotations, and the Appendix, as was done for the articles in this research, this article had the following readability statistics: words per sentence = 23.8, characters per word = 5.3, readability grade = 14.8, Flesch readability level = 42.1, and Kincaid grade level = 13.3. For comparison, this readability grade score would have ranked about 32nd of the 162 articles in our analysis.

APPENDIX
Formulas for Readability Indexes

Kincaid grade level[a] = $11.8 \times$ syllables/words + $.39 \times$ words/sentences − 15.59
ARI[a] = $4.71 \times$ characters/words + $.5 \times$ words/sentences − 21.43
Fog grade level[a] = $.4 \times$ {words/sentence + $100 \times$ [(words $\geq$ 3 syllables)/words]}
SMOG grade level[a] = square root of {[(words $\geq$ 3 syllables)/sentences] $\times$ 30} + 3
Lix[b] = words/sentence + $100 \times$ [(words $\geq$ 6 characters)/words]
Flesch formula[b, c] = 206.835 − $84.6 \times$ syllables/words − $1.015 \times$ words/sentences
Average grade scale = (Kincaid + ARI + Fog + SMOG)/4

[a]One of four scores that are designed to provide a grade at which the material should be readable. Lower scores (grades) suggest more readable material.
[b]Scores varying between 1 and 100.
[c]Flesch is the only readability index for which higher scores indicate better scores.

REFERENCES

AMA (2005), "Evaluation Criteria for Manuscripts," (accessed on November 1, 2005), [available at http://www.marketingjournals.org/jm/ms_evaluate.php].

Armstrong, J. Scott (1980), "Unintelligible Management Research and Academic Prestige," *Interfaces*, 10 (April), 80–86.

——— (1989), "Readability and Prestige in Scientific Journals: Comment," *Journal of Information Science*, 15 (April), 123–24.

Bauerly, Ronald J. and Don T. Johnson (2005), "An Evaluation of Journals Used in Doctoral Marketing Programs," *Journal of the Academy of Marketing Science*, 33 (July), 313–29.

———, ———, and Mandeep Singh (2005), "Readability and the Impact of Marketing," in "Marketing Renaissance: Opportunities and Imperatives for Improving Marketing Thought," *Journal of Marketing*, 69 (October), 19–20.

———, ———, and ——— (2006), "Readability and Writing Well," *Marketing Management Journal*, 16 (1), 216–27.

Baumgartner, Hans and Rik Pieters (2003), "The Structural Influence of Marketing Journals: A Citation Analysis of the Discipline and Its Subareas Over Time," *Journal of Marketing*, 57 (April), 123–39.

Cohen, Jacob (1988), *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ. Lawrence Erlbaum Associates.

Cook, Claire Kehrwald (1985), *Line by Line: How to Edit Your Own Writing*. Boston: Houghton Mifflin.

Crosier, Keith (2004), "How Effectively Do Marketing Journals Transfer Useful Learning from Scholars to Practitioners?" *Marketing Intelligence and Planning*, 22 (May), 540–56.

Dale, Edgar and Jeanne Chall (1949), "The Concept of Readability," *Elementary English*, 26, 19–26.

Desiraju, Ramarao and Steven M. Shugan (1999), "Strategic Service Pricing and Yield Management," *Journal of Marketing*, 63 (January), 44–56.

Dubay, William H. (2004), *The Principles of Readability*. Costa Mesa, CA: Impact Information, (accessed March 15, 2006), [available at http://www.nald.ca/fulltext/readab/readab.pdf].

Flesch, Rudolf and A.H. Lass (1996), *The Classic Guide to Better Writing: Step-by-Step Techniques and Exercises to Write Simply, Clearly and Correctly*. New York: HarperCollins.

Gencturk, Esra F. and Masaaki Kotabe (2001), "The Effect of Export Assistance Program Usage on Export Performance: A Contingency Explanation," *Journal of International Marketing*, 9 (2), 51–72.

Glazer, Rashi (1991), "Marketing in an Information-Intensive Environment: Strategic Implications of Knowledge as an Asset," *Journal of Marketing*, 55 (October), 1–19.

Golder, Peter N. and Gerard J. Tellis (1993), "Pioneer Advantage: Marketing Logic or Marketing Legend?" *Journal of Marketing Research*, 70 (May), 158–70.

Gopen, George D. (2004), *The Sense of Structure: Writing from the Readers Perspective*. New York: Pearson Longman.

——— and Judith A. Swan (1990), "The Science of Scientific Writing," *American Scientist*, 78 (November–December), 550–58.

Hartley, James, James Pennebaker, and Claire Fox (2003), "Using New Technology to Assess the Academic Writing Styles of Male and Female Pairs and Individuals," *Journal of Technical Writing and Communication*, 33 (3), 243–61.

———, Eric Sotto, and James Pennebaker (2002), "Style and Substance in Psychology: Are Influential Articles More Readable Than Less Influential Ones?" *Social Studies of Science*, 32 (April), 321–34.

———, Mark Trueman, and A.J. Meadows (1988), "Readability and Prestige in Scientific Journals," *Journal of Information Science*, 14 (April), 69–75.

Hoch, Stephen J., Xavier Dreze, and Mary E. Purk (1994), "EDLP, Hi-Lo, and Margin Arithmetic," *Journal of Marketing*, 58 (October), 16–27.

Hock, Roger R. (1992), *Forty Studies That Changed Psychology: Explorations into the History of Psychological Research*. Englewood Cliffs, NJ: Prentice Hall.

Holbrook, Morris B. (1986), "A Note on Sadomasochism in the Review Process: I Hate When That Happens," *Journal of Marketing*, 50 (July), 104–108.

Hult, G Tomas M., William T. Neese, and R. Edward Bashaw (1997), "Faculty Perceptions of Marketing Journals," *Journal of Marketing Education*, 19 (Spring), 37–52.

Kintsch, Walter and James R. Miller (1984), "Readability: A View from Cognitive Psychology," in *Understanding Reading Comprehension*, James Flood, ed. Newark, DE: International Reading Association, 220–32.

Klare, George R. (1963), *The Measurement of Readability*. Ames: Iowa State University Press.

Knight, Jonathan (2003), "Scientific Literacy: Clear as Mud," *Nature*, 423 (May), 376–78.

Lutz, Richard J. (1990), "Editorial," *Journal of Consumer Research*, 17 (December), 242–45.

Medoff, Marshall H. (2003), "Article Placement and Market Signaling," *Applied Economics Letters*, 10 (June), 479–82.

Metoyer-Duran, Cheryl (1993), "The Readability of Published, Accepted, and Rejected Papers Appearing in *College and Research Libraries*," *College and Research Libraries*, 54 (November), 517–26.

Mick, David (2005), "Inklings: From Mind to Page in Consumer Research," *Association for Consumer Research Newsletter*, 1 (Summer), 1–3.

Morgan, Neil A., Anna Kaleka, and Constantine S. Katsikeas (2004), "Antecedents of Export Venture Performance: A Theoretical Model and Empirical Assessment," *Journal of Marketing*, 68 (January), 90–108.

Muthén, Linda K. and Bengt O. Muthén (2004), *Mplus User's Guide*. Los Angeles: Muthén & Muthén.

Oppenheimer, Daniel M. (2005), "Consequences of Erudite Vernacular Utilized Irrespective of Necessity: Problems with Using Long Words Needlessly," *Applied Cognitive Psychology*, 20 (October), 139–56.

Peter, J. Paul and Jerry C. Olson (1983), "Is Science Marketing?" *Journal of Marketing*, 47 (Fall), 111–25.

Prasad, V. Kanti, Keshavamurthy Ramamurthy, and G.M. Naidu (2001), "The Influence of Internet-Marketing Integration on Marketing Competencies and Export Performance," *Journal of International Marketing*, 9 (4), 82–110.

Redish, Janice C. and Jack Selzer (1985), "The Place of Readability Formulas in Technical Communication," *Technical Communication*, 32 (August), 46–52.

Robinson, William T. and Sungwook Min (2002), "Is the First to Market the First to Fail? Empirical Evidence for Industrial Goods Businesses," *Journal of Marketing Research*, 39 (February), 120–28.

Ross-Larson, Bruce (1996), *Edit Yourself*. New York: Norton.

——— (1999), *Stunning Sentences*. New York: Norton.

Sawyer, Alan G. (1988), "How to Write a Painful Marketing Manuscript: Prose and Cons," *Journal of Marketing Education*, 10 (Fall), 49–53.

——— and A. Dwayne Ball (1981), "Statistical Power and Effect Size in Marketing Research," *Journal of Marketing Research*, 18 (August), 275–90.

Stewart, David W. (2002), "Getting Published: Reflections of an Old Editor," *Journal of Marketing*, 66 (October), 1–6.

Stremersch, Stefan and Gerard J. Tellis (2002), "Strategic Bundling of Products and Prices: A New Synthesis for Marketing," *Journal of Marketing*, 66 (January), 55–72.

———, Isabel Verniers, and Peter C. Verhoef (2007), "The Quest for Citations: Drivers of Article Impact," *Journal of Marketing*, 71 (July), 171–94.

Strunk, William, Jr., and Elwyn B. White (2000), *The Elements of Style*. Needham Heights, MA: Allyn and Bacon.

Taylor, David and Intuitive Systems (2006), "Readability.info," (accessed May 7, 2006), [available at http://www.readability.info/uploadfile.shtml].

Winer, Russell S. (1998), "From the Editor," *Journal of Marketing Research*, 35 (February), iii–v.

Vorhies, Douglas W. and Neil A. Morgan (2003), "A Configuration Theory Assessment of Marketing Organization Fit with Business Strategy and Its Relationship with Marketing Performance," *Journal of Marketing*, 67 (January), 100–115.

Zinsser, William (1988), *On Writing Well: An Informal Guide to Writing Nonfiction*. New York: Harper and Row.

U.S. Postal Service Statement of Ownership, Management, and Circulation (Required by 39 U.S.C. 3685)

1. Title of Publication *Journal of Marketing*
2. Publication No. 0022-2429
3. Filing Date: October 1, 2007
4. Issue Frequency: Quarterly
5. No. of Issues Published Annually. 4
6. Annual Subscription Price: $95
7. Complete Mailing Address of Known Office of Publication: 311 South Wacker Dr., Suite 5800, Chicago, IL 60606-2266
8. Complete Mailing Address of the Headquarters or General Business Offices of the Publisher 311 South Wacker Dr , Suite 5800, Chicago, IL 60606-2266
9. Full Names and Complete Mailing Address of Publisher, Editor, and Managing Editor: Publisher Jack Hollfelder, 311 S Wacker Dr , Ste 5800, Chicago, IL 60606-2266, Editor Roland T Rust, Dept. of Mktg , Robert H Smith Schl. of Business, 3451 Van Munching Hall, Univ of Maryland, College Park, MD 20742-1815, Managing Editor Francesca V. Cooley, 311 S Wacker Dr , Suite 5800, Chicago, IL 60606-2266
10. Owner American Marketing Association, 311 South Wacker Dr , Suite 5800, Chicago, IL 60606-2266
11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1% or More of Total Amount of Bonds, Mortgages, or Other Securities (If there are none, so state) None
12. For completion by nonprofit organizations authorized to mail at special rates The purpose, function,and nonprofit status of this organization and the exempt status for federal income tax purposes. Has Not Changed During Preceding 12 Months
13. Publication Name: *Journal of Marketing*
14. Issue Date for Circulation Data Below. October, 2007
15. Extent and Nature of Circulation.

	Average no. copies each issue during preceding 12 months	**Actual no. copies of single issue published nearest to filing date**
a. Total No Copies (Net Press Run)	9,300	9,000
b. Paid and/or Requested Circulation		
(1) Paid/Requested Outside-County Mail Subscriptions	8,439	8,195
(2) Paid/Requested In-County Mail Subscriptions	0	0
(3) Sales Through Dealers and Carriers, Street Vendors, Counter Sales & Other Non-USPS Paid Distribution	0	0
(4) Other Classes Mailed Through the USPS	0	0
c. Total Paid and/or Requested Circulation *[Sum of 15b (1), (2), (3) and (4)]*	8,439	8,195
d. Free Distribution by Mail *(Samples, Complimentary, and Other Free) [Sum of 15d (1), (2) and (3)]*	11	11
e. Free Distribution Outside the Mail (Carriers or Other Means)	0	0
f. Total Free Distribution *(Sum of 15d and 15e)*	11	11
g. Total Distribution *(Sum of 15c and 15f)*	8,450	8,206
h. Copies Not Distributed	850	794
I. Total *(Sum of 15g and h)*	9,300	9,000
J. Percent Paid and/or Requested Circulation *(15c/15g x 100)*	99.9% paid circulation	99 9% paid circulation

16. This Statement of Ownership will be printed in the January, 2008, issue of this publication.
17. Signature and title of editor, publisher, business manager, or owner: Jack Hollfelder, Publisher. I certify that the information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions *(including fines and imprisonment) and/or civil sanctions (including multiple damages and civil penalties).*

Rajesh Sethi & Zafar Iqbal

Stage-Gate Controls, Learning Failure, and Adverse Effect on Novel New Products

This article argues that Stage-Gate controls have the potential of restricting learning in a new product development project and thus hurting the performance of novel new products. Specifically, the authors examine whether control on new product development exercised through rigorous gate review criteria increases project inflexibility, which in turn leads to increased failure to learn. They also focus on whether the effect of project inflexibility is worsened when there is turbulence in the firm's technological and market environment. Furthermore, the authors study whether failure to learn can have an adverse effect on new product performance when the product is novel. Finally, they examine whether a certain relaxation in gate evaluation (gate conditionality) that has been suggested recently mitigates the adverse effect of rigorously enforced controls. The results, which are based on a survey of 120 projects that used the Stage-Gate process for new product development, show that repeated application of strictly enforced and objective evaluation criteria for improved control makes projects more inflexible. Gate conditionality does not mitigate the adverse effect of gate review criteria. Project inflexibility leads to learning failure, and this effect is worsened when the technological environment of the firm is turbulent. In turn, learning failure adversely affects the market performance of novel new products.

Keywords: Stage-Gate controls, learning failure, project inflexibility, process improvement, product development teams

During the past two decades, there has been a strong desire on the part of senior managers to control the new product development process in their firms. Such control has been considered important to bring discipline to "chaotic" new product development activity and to manage the process for improved new products, enhanced efficiency, and faster introduction of new products. This desire to control new product development is consistent with the continuing trend in firms toward improving efficiency and lowering costs through process management and control initiatives, such as reengineering, Six Sigma, ISO 9000, and total quality management. A system that offers a methodology for exerting control on new product development is the Stage-Gate process, and consequently it has been widely embraced by firms (Cooper 1998, 2001).[1]

It has been reported that control imposed through rigorous Stage-Gate evaluation helps firms' product development efforts in terms of introducing discipline, improving performance, enhancing efficiency, and reducing new product cycle time (Cooper 1994, 2001; Cooper, Edgett, and Kleinschmidt 2002a). However, there is a growing concern among new product researchers that gate controls are not suitable for all types of products. For example, it has been suggested that the Stage-Gate process is inappropriate for products that are radical innovations (e.g., computerized tomography), because such innovations do not follow the traditional product development process (Leifer et al. 2000; McDermott and O'Connor 2002; Veryzer 1998). An important question is, Are there any problems with the use of the Stage-Gate process for a majority of products that follow the traditional product development process?

In this research, we contend that even in the case of products that are developed using the traditional product development process, rigorous gate controls can harm new products that have high novelty. New product novelty refers to the degree to which a product is unique or distinct from competing alternatives or preceding product generations; for example, transitioning from traditional wired telephones to cordless phones is more novel than transitioning from rotary phones to push button phones (Andrews and Smith 1996; Im and Workman 2004; Sethi, Smith, and Park 2001). Firms actively seek novelty in their new products because such novelty leads to better sales and profits. If rigorous gate controls are indeed not suitable for novel products, by using such controls, firms may be inadvertently compromising the performance of their novel products. However, little is known about why and how rigorous gate controls cause problems for novel new products.

We argue that when gate controls are made rigorous, they have the potential of adversely affecting learning in the new product development project. Such learning is considered critical for effective new product development (Brown and Eisenhardt 1995; Hurley and Hult 1998; Nonaka and

---

[1] Stage-Gate is a registered trademark of Product Development Institute Inc.

Rajesh Sethi is Associate Professor, School of Business, Clarkson University (e-mail: sethir@clarkson.edu). Zafar Iqbal is Associate Professor of Marketing, Kellstadt Graduate School of Business, Department of Marketing, DePaul University (e-mail: ziqbal@depaul.edu). The authors are thankful to Larry Compeau, Danneels Erwin, Larry Feick, Sue Fogel, Stephen Koernig, Somendra Pant, Anju Sethi, and the three anonymous *JM* reviewers for their helpful comments.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (January 2008), 118–134

Takeuchi 1995). To examine how gate controls adversely influence learning and thus harm novel new products, we draw on a variety of literature, such as process management (e.g., Benner and Tushman 2003; Sitkin, Sutcliffe, and Schroeder 1994), management control systems (e.g., Jaworski 1988; Ramaswami 1996), and organizational learning (e.g., Miner, Bassoff, and Moorman 2001; Slater and Narver 1995).

On the basis of these streams of literature, we contend that when gate evaluation becomes rigorous, it has the potential of reducing the flexibility needed for new product development, which in turn can adversely affect learning. Thus, we introduce the concept of project inflexibility, which refers to the extent to which project parameters become rigid and unchangeable when the project is approved after review at initial gates. We develop and test a conceptual link between the extent of rigor in Stage-Gate evaluation criteria and project inflexibility. Then, we examine whether this inflexibility increases postapproval learning failure in the project (i.e., the inability to acquire new information and incorporate it successfully into the project after its approval at the initial stages). In addition, we posit that learning becomes more constrained when the firm's external environment is volatile. Therefore, we also study whether turbulence in the market and technological environment worsens the relationship between inflexibility and postapproval learning failure.

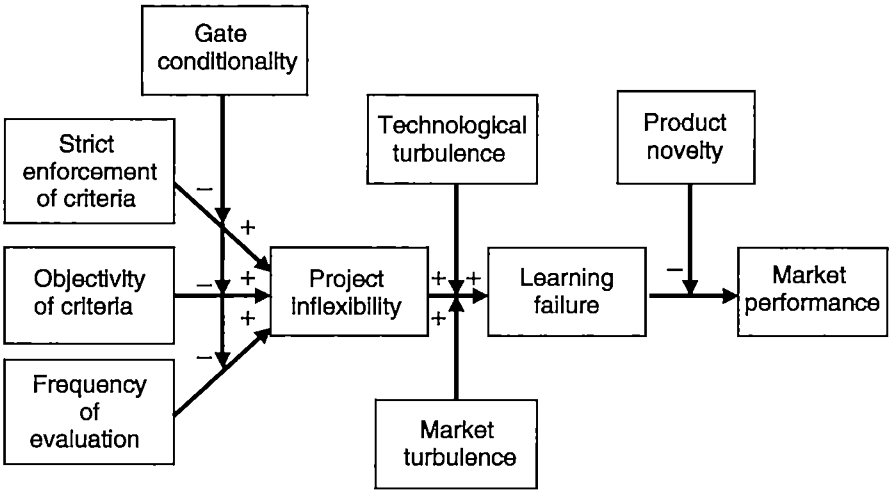
Rigorous Stage-Gate controls may have an adverse effect on project flexibility and learning, but do they harm all types of product development projects equally? The need for learning varies across different kinds of new product development projects (Christensen 1998; Levinthal and March 1993; March 1991). When the product is more unique or novel, learning becomes more crucial. In other words, when there is failure to learn, more novel products may perform poorly compared with less novel products. Thus, we examine the effect of interaction between product novelty and postapproval learning failure on the new product's performance in the marketplace.

If rigorous Stage-Gate controls indeed have an adverse effect on project flexibility and, thus, on learning and the performance of novel products, it is important to ask whether there is a way to mitigate their adverse effect on flexibility. Recently, the Stage-Gate literature has recommended that the problems with rigorous controls can be addressed if gates are made more conditional—that is, if the project is allowed to proceed to the next stage even if it meets just part of the gate criteria, subject to its meeting the remaining criteria at the next or a subsequent gate (Cooper 1994). However, it is important to ask whether simply postponing the application of certain aspects of the criteria through gate conditionality works. As such, this article also studies whether gate conditionality helps reduce the adverse effect of rigorous gate review criteria on project flexibility.

To summarize, this article studies whether Stage-Gate evaluation adversely affects novel new products by examining (1) the effect of the degree of rigor in gate review criteria on project inflexibility, (2) the mitigating influence of gate conditionality on the adverse effect of rigorous gate criteria, (3) the effect of project inflexibility on postapproval learning failure, (4) the moderating influence of market and technological turbulence on the effect of inflexibility, and (5) the effect of learning failure on the new product's performance for different levels of product novelty (for a diagrammatic representation, see Figure 1). Our unit of analysis is a new product development project.

This article makes several contributions. First, we draw on process management and control systems research to provide a conceptual foundation to the Stage-Gate literature. To date, the Stage-Gate literature has been mainly practice driven and lacks a conceptual base. Such a conceptual base is necessary for the systematic development of knowledge for this important and widely used control mechanism. Second, we contribute to the Stage-Gate literature by highlighting that rigorous gate evaluation can harm new products, particularly in the case of novel products and turbulent environments. The main reservation about the utility of the Stage-Gate process has been that these gates can-

FIGURE 1
Hypothesized Model



not be used for game-changing radical innovations, because such innovations use a nontraditional product development process to which gate evaluation cannot be applied. However, the potential adverse effect of such controls on learning and the performance of novel products, particularly in turbulent environments that we study herein, has been overlooked. Third, this study extends the control systems literature by highlighting a trade-off between control and learning. The control literature has not examined whether controls can undercut the effectiveness of an organization by impairing its ability to learn, which can hurt its long-term competitiveness.

Stage-Gate Evaluation, Management Controls, and Process Improvement

Typical Stage-Gate controls break the traditional new product development process into a set of discrete and identifiable stages (e.g., preliminary investigation including concept development and testing, business plan development, product development, testing and validation, full production, and market launch stages), with each stage consisting of a set of prescribed activities. These stages are separated by gates, which serve as the control and go/no-go check points (Cooper 2001). Gates are designed in the form of meetings that take place between senior management and representatives of the product development team. The objective at these gates is to allow projects that adequately satisfy screening criteria to proceed further, thus effectively "killing" bad projects or, at best, allowing bad projects to "recycle" through the previous stages to give them an opportunity to improve (Cooper 2001).

To examine the effect of Stage-Gate controls, it is important to understand the underlying concepts that apply to this process. Gate evaluation basically involves an attempt both to introduce management controls and to improve the process of product development (Cooper 2001). Thus, we discuss the relevant concepts from management control systems and process improvement literature.

Control Systems and Stage-Gate Evaluation

There are two types of formal management controls: process and output controls (Jaworski 1988; Merchant 1985; Ramaswami 1996). Process controls direct by specifying methods and procedures to be adopted for performing tasks. Output controls direct by specifying output goals and standards. The Stage-Gate evaluation process is a combination of both types of management controls. In terms of process control, gate evaluation specifies the steps or the procedure the team must adhere to before each gate. For example, before the project is given a go-ahead for business plan development, the team must develop the product concept with detailed specifications, test the concept on consumers in the target segment, and ascertain consumers' liking for the product idea and willingness to buy. In terms of output control, the project must meet performance standards laid down for the stage. For example, before the project is given a go-ahead for market launch, the test market must show that the product has achieved a certain prespecified level of trial and repeat purchase.

Control systems vary in terms of the degree of rigor. In management control systems, rigor can come from tightness of controls, uniformity or objectivity of controls, and frequency of control (Bechtold 1999; Merchant 1985, 1988). Similarly, rigor in gate criteria can be influenced by how strictly, how objectively and consistently, and how frequently criteria are applied (Cooper 1998, 2001). These three characteristics of rigorous gate criteria need further explanation.

It is argued that evaluation criteria that are not formally applied and strictly adhered to are likely to lead to confusion, ambiguity, and poor evaluation of projects (Cooper 1994, 2001). This is consistent with the recommendation that the gates need to be infused with discipline or "toughened" to enhance the performance of the development process (Cooper, Edgett, and Kleinschmidt 2002a, b). Then, firms are advised to have objective criteria that are not open to interpretation and are consistently applied across various projects (Cooper 1998). Gate evaluation criteria are usually operationalized in the form of standard checklists or scorecards for better prioritization decisions (Cooper 2001; Cooper, Edgett, and Kleinschmidt 2002a, b; Griffin 1997). This requirement of objectivity is meant to ensure that all projects are evaluated against the same criteria, so that different project managers and gate evaluators do not interpret the criteria as it suits them to get their favored projects approved at the gates. When the same gate criteria are consistently applied across all projects and no variations are allowed for the nature of the project, such criteria are likely to be viewed as restrictive and rigorous. Finally, it is suggested that the project should be evaluated against the criteria at various stages throughout the development process. Typical Stage-Gate processes have between four and seven gates. Indeed, practitioners have been recently advised to formally add additional stages and gates to incorporate activities such as ideation and discovery (Cooper 2001; Cooper, Edgett, and Kleinschmidt 2002a, b). Although such frequent evaluation can become restrictive, it is considered important to ensure that the project continues to adhere to the criteria and agreed-on targets and not go offtrack. We focus on strict enforcement, objectivity, and frequent application of the evaluation criteria as three independent variables in our study.

Some of these guidelines for designing gate controls tend to play a role in making such controls rigorous. In addition, the way these gate controls are actually implemented in companies can increase their rigor, even though some literature cautions that the Stage-Gate process should not be made very restrictive (Cooper 1994, 2001). In their desire to exert firm control, managers can sometimes make the evaluation process overly rigorous (Bechtold 1999; Merchant 1985).

Furthermore, as the literature suggests, gate controls can be specific in nature. Some examples of the criteria at each of the gates that appear in the literature are market potential (at the idea-screening gate), customer acceptance (at the concept-screening gate), expected sales in units (at the business analysis gate), product performance or quality (at the product-testing gate), and customer satisfaction (at the test market gate) (Carbonell, Escudero, and Aleman 2004; Hart et al. 2003). Some other examples of specific criteria firms use are return on investment (ROI), concept test scores, test market sales, prototyping costs, and time to development (Cooper 2001).

Among other things, the literature on management control systems has focused on employees' negative reaction to control systems, such as data manipulation, invalid data reporting, and working against the interest of the company (Jaworski 1988; Merchant 1985; Ramaswami 1996). However, in this article, we examine a problem that has not received attention in the management control systems literature—namely, the adverse effect of Stage-Gate evaluation on learning in a new product development project. To do so, we need to discuss some of the relevant concepts from the process management literature.

Process Management and Stage-Gate Evaluation

Process management is a core component of several initiatives, including ISO 9000, Business Process Reengineering, Six Sigma, and total quality management programs. Processes are collections of activities that, taken together, produce output for customers (Benner and Tushman 2003; Garvin 1995). These customers can be either external or internal customers who are recipients of output from an upstream process in the organization. Although process management techniques were first used in the domain of manufacturing and operations improvement, the influence of process management has gradually spread into product innovation. More recent work has focused on how process management affects innovation (Benner and Tushman 2003; Garvin 1995).

The focus of process management techniques is on improving efficiency and reducing variation (Benner and Tushman 2002; Garvin 1995). Although the initial process management literature is more prescriptive in nature, some emerging literature has begun to develop theoretical explanations for how process management techniques work to achieve their objectives (e.g., Benner and Tushman 2003). Among other things, this literature suggests that process management approaches end up stabilizing organizational routines and tightening linkages among them (Benner and Tushman 2002, 2003). Thus, as process management techniques take root in organizations, they tend to reduce adaptability or flexibility (Benner and Tushman 2003). Because the spirit of the Stage-Gate process is similar to that of other process management initiatives, we examine the role of gate evaluation in increasing project inflexibility. Formally, project inflexibility can be defined as the extent to which project parameters (e.g., product definition, concept, specifications, architecture) are rigid and unchangeable when the project is approved after review at the initial gates. In other words, our specific focus is on postapproval project inflexibility. The importance of flexibility for effective new product development has been highlighted in the innovation literature (Iansiti and MacCormack 1997; MacCormack, Verganti, and Iansiti 2001; Moorman and Miner 1998).

Theory and Hypotheses

Effect of Gate Review Criteria on Project Inflexibility

In keeping with process management philosophy, the Stage-Gate process has aims such as streamlining the new product development activity, ensuring that all new product development projects pass through standard product development stages, and making the decision-making and resource allocation processes more uniform for various projects. As process management techniques lead to more efficiency, reductions in costs, and improved speed, the organization pursues these techniques even more enthusiastically, and gradually, variation reduction methods become more stabilized and deeply embedded in the organization's routines (Benner and Tushman 2002, 2003). Similarly, when an organization embraces the Stage-Gate process, it is also likely to diminish costs and wastage due to the elimination of weaker projects, improved speed to the market, and a more equitable and efficient way of allocating scarce resources to various projects (Cooper 1994; Mankin 2004). Because these results help improve the bottom line, the organization is likely to have enhanced faith in the Stage-Gate process (Leonard 1992; Mankin 2004). Thus, gate evaluation becomes more embedded in new product development routines and is rigorously implemented.

In an organizational climate of diminishing resources, product development teams are concerned about their project's continued survival (Dougherty and Heller 1994; Kanter 1983). If gate meetings are situations in which the future of the development effort could be in jeopardy, it becomes imperative to ensure that the strictly enforced criteria are met. For example, in response to gate requirements, product development team members may strive to define the project and product specifications in detail early in the development process, even if they believe that it is too early to freeze such details (Krishnan and Bhattacharya 2002). When precise project parameters are presented to the gate review committee to win its approval, a commitment usually occurs between the product development team and the committee that the team will design the product with these parameters (MacCormack, Verganti, and Iansiti 2001; Verganti 1999). The committee may use the agreed-on parameters as the standard against which it will evaluate the progress of the project at subsequent gates. If there is strict enforcement of gate criteria, the committee can view any changes in the parameters as diluting the project or disturbing the standard.

Similarly, strict criteria may require that projects provide ROI estimates early in the evaluation process. Unless the project team can convince the committee that the project will produce some minimum ROI, the project may not be funded. To estimate the ROI, the team must figure out the investment needed for the project, and for that to happen, it must, for example, finalize the technology and manufacturing facility that will be used. After the team makes these choices and presents them to the committee as the basis for estimating the ROI, in a strict enforcement situation, the team may not be able to change its decisions easily.

When the criteria are highly objective and consistently applied across all projects, they are also likely to impose constraints on the project. For example, a project team may be considering using some technology with which it has little experience. Thus, it may not be possible to foresee what form the final product will adopt. The criteria that do not allow different interpretation from project to project are likely to require the project to follow the same sequence of activities and meet the same criteria that are expected from a more typical project. Thus, the project team may need to make premature choices that may not be easily reversible (MacCormack, Verganti, and Iansiti 2001).

Finally, although frequent gate evaluation can help the product development team by providing repeated feedback on its progress (Bechtold 1999; Merchant 1985), it can harm project flexibility. With frequent evaluation, the team ends up committing to deliverables repeatedly and, thus, more deeply to particular specifications and targets. Then, the team must frequently show that it is fulfilling the required criteria and is adhering to all its previously agreed-on targets and commitments. Thus, with repeated evaluations, the project parameters are expected to become more difficult to change.

In summary, with more strictly enforced, objective, and frequently applied gate review criteria, the project is more likely to lose its flexibility. Thus:

H_1: The more strictly the gate review criteria are enforced, the greater is the inflexibility of the new product development project.

H_2: The more objective the gate review criteria, the greater is the inflexibility of the new product development project.

H_3: The more frequent the evaluation of the project at gates, the greater is the inflexibility of the new product development project.

Moderating Influence of Gate Conditionality

Gate conditionality means that projects are allowed to proceed further into the process of development conditional on their meeting required criteria at a subsequent stage or that certain project activities are approved out of the usual sequence (Cooper 1994). Although all criteria must still be met for a product to be finally launched, gate conditionality implies that a project will not be held up because of certain criteria remaining unmet at a particular time. As long as the criteria are met at a subsequent gate, the project is allowed to proceed further, albeit conditionally. In other words, the restrictive go/no-go nature of gates is somewhat relaxed.

Making gates conditional is consistent with the concept of relaxing tight controls in the control systems literature (Bechtold 1999; Merchant 1985). In some situations, granting some leeway can help minimize the adverse effect of tight controls. Basically, if the product team can get a project conditionally approved and can meet some of the mandatory criteria at a subsequent gate, despite the presence of strict controls, it has more freedom to make changes in the project (Cooper 1994). Even if the criteria are objectively and consistently applied and are strictly enforced, the freedom the team has to meet them at a subsequent gate can save it from making premature choices about technology and customer needs. Similarly, although the frequency of evaluation may be high, because the team can get conditional approval frequently, it can explore alternative approaches to develop better specifications and will not need to commit itself repeatedly to meeting project parameters with which it is not comfortable. Thus:

H_4: The greater the gate conditionality, the weaker is the impact of strictly enforced gate review criteria on the inflexibility of the new product development project.

H_5: The greater the gate conditionality, the weaker is the impact of objective gate review criteria on the inflexibility of the new product development project.

H_6: The greater the gate conditionality, the weaker is the impact of frequent gate evaluations on the inflexibility of the new product development project.

Effect of Project Inflexibility on Postapproval Learning Failure

Organizational learning involves acquiring and processing new information continuously to elevate knowledge and improve decisions and actions based on the new knowledge (Fiol and Lyles 1985; Miner, Bassoff, and Moorman 2001; Slater and Narver 1995). New product development entails learning about markets and technologies and matching them to create offerings that satisfy targeted customers. The new product development process within firms needs to incorporate mechanisms for learning, such as market assessment, technology assessment, product testing, prototype testing, test marketing, and launch assessment (Cooper 2001; Crawford and DiBenedetto 2002). The ability to learn continuously enables firms to stay abreast of changes in markets and technologies.

How does project inflexibility lead to postapproval learning failure? Learning failure refers to a team's inability to acquire and process any new information, integrate the new information into the project, and make changes in plans when the project is approved after initial gates. This conceptualization of learning failure is based on the organizational learning literature that equates learning not just with acquiring information but also with incorporating that information into conceptualizations and plans and changing action repertoire (Duncan and Weiss 1979; Fiol and Lyles 1985; Hedberg 1981).

When a project becomes inflexible after approval at gate reviews, it can restrict the product development team. The team has reduced freedom to change the product definition and specifications (MacCormack, Verganti, and Iansiti 2001). Then, the team can lose the freedom to use or divert resources in a manner it deems fit. For example, if the team has not sought approval at prior gates for using a particular technology for the product and subsequently feels the need to do so, it may not be easy to divert resources meant for some other purpose to explore the suitability of this technology. Thus, the project's high inflexibility can increase the team's level of frustration.

Such restrictions can have an adverse effect on a team's ability to learn. According to the organizational learning literature, effective learning takes place if there is a perceived knowledge gap, motivation to acquire new knowledge, and cognitive freedom to process new input and change the existing cognitive structure (Duncan and Weiss 1979; Fiol and Lyles 1985; Hedberg 1981). If the project parameters are rigidly fixed, the team is expected to overlook any knowledge gaps and may have reduced motivation to acquire any new information or take a fresh look at the existing information. In other words, not only are team members likely to stop actively searching for market- or technology-related information, they may ignore some input even if it becomes available to them. Furthermore, team members' struggle and frustration with stringent project parameters can adversely affect their cognitive flexibility to accommodate any new input or to alter their existing knowledge structure and accordingly make changes in plans (Canas, Fajardo, and Salmeron 2006; Duncan and Weiss 1979). In effect, learning failure occurs in the team. Thus:

> H_7: The higher the project inflexibility, the greater is the postapproval learning failure.

Moderating Effect of Turbulence in the Market and Technological Environment

Market turbulence is defined as the degree of instability, uncertainty, and lack of control within a firm's marketplace (Jaworski and Kohli 1993). Technological turbulence refers to the degree to which technology changes over time within an industry and the degree to which such changes affect the industry (Jaworski and Kohli 1993). When there is volatility in the market and technological environment, new information about customers, competitors, and the technology can unexpectedly emerge from time to time. Although team members would like to take such new information into account for the benefit of the project, if project inflexibility is high, they are faced with a situation in which they cannot make major changes to the project after its approval. As we discussed previously, when project parameters become rigid, team members' motivation to search for information and cognitive flexibility to make changes in their knowledge structure tend to be adversely affected. However, turbulence in the market and technological environment creates many unexpected developments that require a good deal of motivation and cognitive flexibility to accommodate them into the existing knowledge structure (Canas, Fajardo, and Salmeron 2006; Duncan and Weiss 1979). This mismatch between low motivation and cognitive flexibility and high need for these two factors in turbulent environments is likely to frustrate team members. Such frustration will worsen members' attention to new information and its integration into project plans. Thus:

> H_8: The greater the turbulence in the technological environment, the stronger is the impact of project inflexibility on postapproval learning failure.
>
> H_9: The greater the turbulence in the market environment, the stronger is the impact of project inflexibility on postapproval learning failure.

A key goal of the Stage-Gate process is to improve the success of new products developed by weeding out ineffective product ideas and plans. However, as we contended, rigorous gate controls can adversely affect postapproval learning in the product development project and thus have the potential of harming the performance of novel new products.

Effect of Postapproval Learning Failure on Market Performance of Novel Products

The literature has highlighted the crucial role of organizational learning in new product development (Brown and Eisenhardt 1995; Diamond and Iqbal 2000; Hurley and Hult 1998; Nonaka and Takeuchi 1995). However, the need for learning is likely to vary depending on the novelty of the product (Atuahene-Gima 2005; Danneels 2002). Novelty has been defined as the extent to which the idea for a new product is unique and original compared with other competing products (Amabile 1983; Andrews and Smith 1996; Im and Workman 2004; Jackson and Messick 1965; Sethi, Smith, and Park 2001).

Typically, products with high novelty involve a great deal of exploratory learning activities on the part of companies (March 1991; Sethi, Smith, and Park 2001). If the product idea is novel, it is often difficult to obtain all the input and knowledge required for it in the initial stages of product development (MacCormack, Verganti, and Iansiti 2001). Thus, learning in the subsequent stages of product development becomes even more critical. In general, if the idea is novel, the product development team is likely to begin the project with some tentative knowledge about the market and technology that needs to be refined and updated at almost every step of the development process. Thus, unless a team can update its knowledge persistently in the subsequent stages of the new product development process, the product it develops may not be appropriate for the market. If the product is novel, postapproval learning failure can harm the market performance of the product. Conversely, if the product lacks novelty, learning is less critical for product success. Therefore, we offer only an interaction effect hypothesis.

> H_{10}: The greater the product novelty, the stronger is the adverse effect of postapproval learning failure on the market performance of the new product.

Methodology

Data Collection Procedure

Before developing a survey to test our hypotheses, we conducted interviews with 20 new product development executives across 12 industries, all of whom had experience with Stage-Gate development processes. In total, we used approximately 25 hours of interviews to gain an understanding of the benefits and problems new product development executives experienced when navigating through the Stage-Gate process in their firms. These interviews also helped develop the measures for the empirical study.

We tested hypotheses using data collected through a survey from key informants involved in recent new product initiatives. Following similar key informant research, our goal was to identify a person who would be highly knowledgeable about the product development project and the Stage-Gate process (Sethi, Smith, and Park 2001). In this study, this person was the manager who was responsible for a new product development project.

The criteria for participation in the survey were that (1) the executive surveyed should have completed a new product development project within two years preceding the survey as a project manager, (2) the Stage-Gate process should have been used for the project, and (3) a team should have been involved in the development of the product. To identify candidate projects and related key informants, we searched secondary sources of information and created a list of firms and their divisions that developed and introduced new products within two years preceding this survey. We contacted the short-listed firms, identified project managers who met the preceding criteria, and requested their participation in our study. We asked the managers who agreed to participate to complete questionnaires that were e-mailed to them. Managers were given the option of either faxing or mailing their surveys back to the authors.

In completing the questionnaire, respondents were asked to focus on their most recently completed new product development project. We sent questionnaires to 200 project managers. We e-mailed a reminder with another copy of the questionnaire to managers who had not responded after two weeks. We received 120 usable responses, for a 60% response rate. The industries in the sample mainly consisted of telecommunications, light engineering products, chemical products, consumer software, food, and home appliances.

On average, projects had been completed ten months before data collection (85% of them were completed in the prior year), thus making them recent enough to ensure reasonably accurate recall of the team experience. To examine the possibility of retrospective bias in the data, we checked whether there were any differences in responses between projects completed more recently and those completed earlier (based on a median split). There were no significant differences in responses in the two categories, suggesting that retrospective bias should not be a cause for concern in our study.

To assess the degree of nonresponse bias, we divided responses into two categories—the responses received before reminders and those received after reminders. We conducted t-tests to examine the possible difference between groups using mean responses to each of the variables included in the model. We made these calculations under the assumption that those who responded late were similar to nonrespondents (Armstrong and Overton 1977). There were no significant differences between groups. As a further check for nonresponse bias, we randomly contacted 12 managers who did not respond to the survey to explore the reasons for nonresponse and to check whether their projects or companies differed in some way from those that responded. These managers told us that they did not respond either because they were busy around the time the survey reached them or because, after examining the survey, they believed that their company would not permit sharing of the information required to be included in the survey. The new product development projects, the Stage-Gate process, and the companies of nonresponding managers were typical of the projects and companies that were in the sample. Therefore, we inferred that managers who responded were not very different from those who did not respond.

On average, in our sample, firms had seven gates in their Stage-Gate processes. The Stage-Gate process had been in place for an average of five years. The executives that completed the survey had an average new product development experience of seven years. The respondents reported that they were knowledgeable about various aspects of the project. The average knowledge on a seven-point scale was 5.92 (7 = "highly knowledgeable"). The respondents had diverse backgrounds: marketing (38%), product development and research and development (31%), general management (15%), and manufacturing (16%). There were no significant differences on study variables among respondents with different backgrounds. In addition, there were no significant differences between the responses by mail and those by fax.

Measures

Whenever possible, we used existing measures of constructs. For the newly developed measures, we created a pool of items based on the existing literature and our interviews with managers. We pretested all measures on 15 project managers and academics. On the basis of the pretest, we refined the measures. We conducted detailed interviews with 5 of these managers to ascertain their interpretation of the items in the measures. This feedback helped us further revise the questionnaire, which we then administered to the full sample.

Because the unit of analysis in this study was a project, several constructs in the study were measured at the project level. Given that all the projects were managed by teams, we considered it equally appropriate to operationalize constructs from the perspective of either the project or the project team. Descriptive statistics and reliability for each refined measure appear in Table 1, and correlations among measures appear in Table 2. The Appendix shows both the final scale items and those that we dropped during scale refinement.

TABLE 1
Descriptive Statistics of Refined Measures

Construct	Number of Items	M[a]	SD[a]	α
Strictly enforced gate review criteria	3	4.31	1.58	.89
Objective gate review criteria	3	3.53	1.26	.74
Frequency of evaluation	1	7.28	4.76	—
Project inflexibility	3	4.04	1.20	.79
Gate conditionality	3	4.74	1.07	.67
Learning failure	3	3.70	1.25	.82
Technological turbulence	4	4.19	1.15	.82
Market turbulence	4	4 64	1.36	.87
Market performance	3	5.86	2.01	—
Product novelty	4	4.64	1.22	.92

[a]All were evaluated on a seven-point scale except for frequency of evaluation.

TABLE 2
Correlations

	1	2	3	4	5	6	7	8	9	10
1. Strictly enforced gate review criteria	1.00									
2. Objective gate review criteria	.53***	1.00								
3. Frequency of evaluation	–.48***	.16*	1.00							
4. Project inflexibility	.53***	.37***	.32***	1.00						
5. Gate conditionality	–.20**	–.40***	–.11	–.19**	1.00					
6. Learning failure	.23**	.11	36***	.44***	–.24***	1.00				
7. Technological turbulence	.18	.04	.34***	.12	–.12	.17	1.00			
8. Market turbulence	.16	05	.27***	.21**	–.08	.24***	.27***	1.00		
9. Market performance	.13	.02	.15	.02	–.08	–.03	.14	–.13	1.00	
10. Product novelty	18*	.09	–.11	.11	–.02	.02	–.05	.06	.01	1.00

*Significant at $p < .10$ (two-tailed).
**Significant at $p < .05$ (two-tailed).
***Significant at $p < .01$ (two-tailed)

Strictly enforced gate review criteria. We developed a new measure for this construct to assess the degree to which the gate review criteria were formally established and used and strictly enforced. The items in the measure were based on the Stage-Gate literature (Cooper 1994, 2001). The scale for strictly enforced criteria is a seven-point, three-item, Likert-type scale ($\alpha = .89$). An example of an item for this measure is, "The criteria for gate reviews are strictly adhered to."

There was sufficient variability in the data on strictness of gate review criteria (SD = 1.58 on a seven-point scale). In the top 25% of the strictness scale, there were 29 projects, and in the bottom 25%, there were 41 projects. This distribution suggests that though several firms have strict gate controls, there are many with lenient controls.

Objective gate review criteria. On the basis of Cooper's (2001) recommendations about Stage-Gate evaluation, we developed a new measure to assess the extent to which projects were reviewed using criteria that were not open to subjective interpretation or negotiation from project to project. An item for this measure is, "The criteria for gate reviews are open to negotiations" (reverse coded). The scale for objective criteria is a seven-point, three-item, Likert-type scale ($\alpha = .74$).

There was sufficient variability in objectivity of Stage-Gate criteria across projects (SD = 1.26 on a seven-point scale). There were almost an equal number of projects at both the top end and the bottom end of the objectivity scale. For example, in the top 25% of the objectivity scale, there were 35 projects, and in the bottom 25%, there were 34 projects.

Frequency of evaluation. We asked the respondents to report the total number of gates in their company's Stage-Gate process as a measure of frequency of evaluation.

Project inflexibility. To develop this measure, we drew on both flexibility literature (MacCormack, Verganti, and Iansiti 2001) and interviews with managers. Specifically, we asked respondents to rate the extent to which project parameters were fixed with little room for change after gate reviews. The project inflexibility scale is a seven-point, three-item, Likert-type scale ($\alpha = .79$). An example of an item for this measure is, "After approval at gate reviews, the approved project parameters were to be adhered to in a fairly rigid manner."

There was a good deal of variance in inflexibility across projects (SD = 1.20 on a seven-point scale). There were several projects at both the top end and the bottom end of the inflexibility scale. For example, 35 projects had a high degree of inflexibility (in the top 25% of the scale), and 28 projects had little inflexibility (in the bottom 25% of the scale).

Gate conditionality. We developed this measure on the basis of our interviews with new product executives and from Cooper's (1994) normative prescriptions. The measure assessed the extent to which the gates allowed projects to be approved even if they did not meet all criteria, conditional on the criteria being met at future gates. We used a seven-point, three-item, Likert-type scale to measure the degree of gate conditionality in the Stage-Gate process ($\alpha = .67$). An item for this measure is, "Gate reviews of projects allow a project to be approved even if all criteria are not met, conditional on those criteria being met in the future."

Learning failure. This measure assessed the degree to which project members focused less on new information and found it difficult to incorporate new information in the project after approval at initial gates. We developed the measure on the basis of the learning literature (Duncan and Weiss 1979; Fiol and Lyles 1985). We used a seven-point, three-item, Likert-type scale to measure learning failure ($\alpha = .82$). An example of an item for this measure is, "After approval, project members had a tendency to focus less on new information (e.g., market- and/or technology-related information) that could have helped improve the project."

Technological turbulence. This measure assessed the extent to which rapid technological changes and breakthroughs occurred in the environment leading to product development opportunities. The measure was based on Jaworski and Kohli's (1993) measure of technological turbulence. We used a seven-point, four-item, Likert-type scale for technological turbulence ($\alpha = .87$). An item for this measure is, "The technology in our industry is changing rapidly."

Market turbulence. We operationalized market turbulence as the degree of instability that existed in the environment with respect to customer needs and competitive actions. The measure was based on a measure that Jaworski and Kohli (1993) use. We used a seven-point, four-item, Likert-type scale for market turbulence ($\alpha = .82$). An example of an item for this scale is, "It is difficult to predict how customers' needs and requirements will evolve in our markets."

Product novelty. We adapted this measure from an established measure that Sethi, Park, and Smith (2001) use. We asked the key informants to rate the new product compared with competing products on a battery of bipolar adjectives, such as novel versus predictable, original versus commonplace, unique versus ordinary, and unusual versus usual. The scale for novelty was a seven-point, four-item, semantic scale ($\alpha = .92$).

New product market performance. We adapted the new product market performance measure from an existing measure used in prior studies (e.g., Sethi 2000). The scale for market performance is a seven-point, three-item, Likert-type scale. We asked the key informants to rate the new product's performance on sales, market share, and profits compared with planned objectives for the product. The measure of performance is a formative measure (Diamantopoulos and Winklhofer 2001).

Covariates and Their Measures

To control for other possible explanations of inflexibility and new product performance, we include several covariates in this study. We now discuss these covariates and their measures.

Covariates affecting project inflexibility. The covariates of interest are the extent to which the project was influenced by competitive considerations, abstractness of the gate criteria, industry (business-to-business versus business-to-consumer), the duration of the project, and the degree of formalization in the firm. If the project is influenced by competitive considerations, its parameters may be driven by competitive products and thus may be less changeable (Robinson and Chaing 2002). We measured the influence of competitive considerations with a seven-point, single-item, Likert-type scale. If the gate criteria are more abstract than specific, the project parameters may not become so rigid. We measured criteria abstractness versus specificity with a seven-point, single-item, Likert-type scale. Because business-to-business products are more likely to be designed in response to defined requirements of industrial customers, we expect that their specifications are more inflexible than products in the business-to-consumer industry (Hutt and Speh 2004). We operationalized industry type as a dummy variable categorizing the products as either consumer products (business-to-consumer) or business products (business-to-business). Furthermore, for projects of much shorter duration, specifications might be frozen soon after the start of the project, thus introducing early inflexibility in the project (Smith and Reinertsen 1991). We measured project duration by asking respondents how many months it took to complete the project. Finally, a high degree of formalization in the firm may have an adverse effect on flexibility (Zaltman, Duncan, and Holbek 1973). We measured formalization using a seven-point, eight-item, Likert Type scale ($\alpha = .72$) that we adapted from the work of Jaworski and Kohli (1993) and Menon and colleagues (1999).

Covariates affecting performance. The covariates of interest are the duration of the project, expenditure on the project, and environmental munificence. Projects of longer duration may provide more time for careful product development and thus can increase the likelihood of success (Crawford 1993). If the firm makes more money available to a project than the standard for comparable projects, it might influence the superiority and success of the product. We assessed the expenditure relative to other similar projects in the firm using a seven-point, single-item, Likert-type scale. Finally, if the firm's environment is munificent (i.e., markets are growing, and higher product margins can be achieved), the new products are more likely to succeed (Dess and Beard 1984). We measured environmental munificence using a seven-point, three-item scale ($\alpha = .82$) that asked the respondents to report the extent to which the firm's markets were growing and the firm enjoyed higher margins.

Scale Testing

The refined scales were the result of exploratory and confirmatory factor analyses. We performed an exploratory factor analyses with principal components extraction and Varimax rotation on all scales used to examine the effect of review criteria on project inflexibility (i.e., strictly enforced criteria, objective criteria, gate conditionality, and project inflexibility). The results showed the anticipated factor structure; items loaded highly on the constructs they were intended to measure.

We performed another exploratory factor analysis on the remaining constructs in the study (i.e., market turbulence, technological turbulence, learning failure, and product novelty). Again, the items loaded highly on their own constructs. In both exploratory analyses, the scales were unidimensional.

We then subjected the measures to confirmatory factor analyses. Similar to exploratory factor analysis, we first assessed a four-factor confirmatory model, including all constructs used to examine the effect of review criteria on project inflexibility (i.e., strictly enforced criteria, objective criteria, gate conditionality, and project inflexibility), through structural equation modeling with LISREL. The results indicated good fit of the confirmatory measurement model of the four constructs ($\chi^2 = 79.71$, $p = .003$; goodness-of-fit index = .91; comparative fit index = .95; and root mean square error of approximation = .075). In the second confirmatory model, we included the remaining constructs—namely, market turbulence, technological turbulence, learning failure, and novelty. Again, the results indicated good fit of the confirmatory measurement model of the four constructs ($\chi^2 = 193.76$, $p = .04$; goodness-of-fit index = .90; comparative fit index = .98; and root mean square error of approximation = .052). In both confirmatory analyses, all items loaded on their respective constructs, and each loading was significant ($p < .001$).

Results

Hypothesis Testing

We tested all the hypotheses (H_1–H_{10}) using moderated regression analysis, as Aiken and West (1991) and Jaccard, Wan, and Turrisi (1990) suggest. To minimize multicollinearity between interaction terms and their constituent terms in the regression model, we mean-centered measures involved in interaction terms (Aiken and West 1991; Jaccard, Wan, and Turrisi 1990). As Table 3 shows, 28% of the variance in project inflexibility is accounted for by independent variables and moderators in the model (for H_1–H_6), and the F-statistic is 4.87 ($p < .001$). Examination of variance inflation factors (VIFs) among the explanatory variables indicates that the highest VIF was 2.84, suggesting that multicollinearity was not a problem (Neter, Wasserman, and Kutner 1990).

Furthermore, project inflexibility and market and technological turbulence account for 20% of the variance in learning failure, and the F-statistic is 6.87 ($p < .001$). The highest VIF among the explanatory variables is 1.13. Finally, learning failure and product novelty (H_8) account for 10% of the variance in product market performance, and the F-statistic is 3.17 ($p < .01$). Even in this model, the highest VIF among the explanatory variables is 1.18, suggesting that there was no problem of multicollinearity.

Effects of Gate Review Criteria on Project Inflexibility

We find that strict gate review criteria are positively related to project inflexibility (b = .21, t = 2.06, $p < .01$), in support of H_1. We predicted that objective gate review criteria would have a positive relationship to project inflexibility. We find that the relationship is significant and positive, in support of H_2 (b = .20, t = 1.72, $p < .05$). H_3, which predicted a positive impact of frequency of evaluation on project inflexibility, is also supported, though the effect is marginal (b = .10, t = 1.40, $p < .10$).

We also wanted to examine whether the gate criteria have a nonlinear relationship to project inflexibility because moderate levels of frequent evaluation, objectivity, and strict enforcement might actually bring some needed discipline and help the project. Therefore, we tested the model with quadratic terms for frequency, objectivity, and strictness. These terms are not significant, suggesting that their effects are linear.

Moderating Effect of Gate Conditionality

We expected gate conditionality to weaken the positive effects of strictly enforced gate review criteria. However, the interaction between gate conditionality and strictly enforced evaluation criteria is not significant (b = .01, t = .07, $p = .94$). Thus, H_4 is not supported. H_5, which predicted a mitigating effect of gate conditionality on the influence of objective criteria on project inflexibility, is also not supported (b = .01, t = .20, $p = .85$). Similarly, H_6, which predicted a weakening effect of gate conditionality on the relationship between frequency of evaluation and project inflexibility, is not supported (b = .01, t = .10, $p = .92$). In other words, the expectations in the literature about the beneficial effect of gate conditionality are not realistic.

We also wanted to explore whether the effect of the gate criteria on project inflexibility is worsened in the case of novel projects. Therefore, we tested the moderating effect of gate novelty. However, the interaction effect is not significant, suggesting that novelty plays no such role.

Covariates for Project Inflexibility

The influence of competitive considerations does not affect project inflexibility (b = –.11, t = –.36, $p = .73$). Similarly, even if the criteria are more abstract and less specific, they do not influence project inflexibility (b = –.12, t = –.36, $p = .72$). In addition, the duration of the project has no effect on project inflexibility (b = –.01, t = –1.26, $p = .21$). Formalization in the firm increases project inflexibility (b = .09, t = 1.81, $p < .05$), but industry has no effect (b = –.39, t = –.57, $p = .56$).

Effect of Project Inflexibility on Postapproval Learning Failure

Consistent with our expectation, project inflexibility is positively related to postapproval learning failure (b = .40, t = 4.37, $p < .01$). Thus, H_7 is supported. Furthermore, technological turbulence worsens the influence of project inflexibility on learning failure (b = .03, t = 1.63, $p < .05$), in support of H_8. However, contrary to our expectation that market turbulence positively moderates the relationship between project inflexibility and learning failure, we find no moderating effect (b = –.01, t = –.45, $p = .33$). Thus, there is no support for H_9.

We also wanted to examine the effect of project inflexibility on the three aspects of learning failure—namely, the inability to acquire new information, to integrate the information into the project, and to make changes in plans accordingly. Therefore, we correlated project inflexibility with the three aspects of learning failure. The results show that project inflexibility adversely affects all three aspects, but some more than others. Specifically, project inflexibility increases failure to acquire new information (r = .49, $p < .01$), failure to incorporate the information into the project (r = .39, $p < .01$), and failure to make changes in plans based on information (r = .21, $p < .01$). Project inflexibility has the worst effect on information acquisition among the three aspects of learning failure and the least adverse effect on making changes based on new information.

We wanted to explore whether product novelty also moderates the relationship between project inflexibility and learning failure. Therefore, we included the interaction term between project inflexibility and product novelty in the model. This term is not significant, suggesting that novelty does not moderate the relationship between project inflexibility and learning failure.

Effect of Learning Failure and Product Novelty on Product Market Performance

We hypothesized that the interaction between learning failure and product novelty would have a negative relationship

TABLE 3
Results

Variables in the Equation	b	β	t-Value
1. Dependent Variable = Project Inflexibility			
Main Effect Variables			
Strictly enforced gate review criteria	.21	.27	2.06***
Objective gate review criteria	.20	.20	1.72**
Frequency of evaluation	.10	.13	1.40*
Interaction Terms			
Strictly enforced review criteria × gate conditionality	.01	.01	.07
Objective review criteria × gate conditionality	.01	.02	.20
Frequency of evaluation × gate conditionality	.01	.02	.10
Moderator Variable			
Gate conditionality	–.08	–.07	–.42
Covariates			
Competitive consideration	–.11	–.05	–.73
Abstract criteria	–.12	–.05	–.36
Duration of project	–.01	–.10	–1.26
Formalization	.09	.15	1.81**
Industry	–.39	–.05	–.57
Adjusted R^2 = .28, F-statistic = 4.87*** (d.f. = 12, 107)			
2. Dependent Variable = Learning Failure			
Main Effect Variable			
Project inflexibility	.41	.38	4.37***
Interaction Terms			
Project inflexibility × technological turbulence	.03	.14	1.63**
Project inflexibility × market turbulence	–.01	–.04	–.45
Moderator Variables			
Technological turbulence	.06	.08	.34
Market turbulence	.10	.12	1.35*
Adjusted R^2 = .20, F-statistic = 6.87*** (d.f. = 5, 114)			
3. Dependent Variable = Market Performance			
Interaction Term			
Learning failure × product novelty	–.04	–.16	–1.70**
Main Effect Variable			
Learning failure	–.09	–.06	–.69
Moderator Variable			
Product novelty	–.05	–.05	–.52
Covariates			
Duration of project	.01	.27	3.10***
Project expenditure	–1.04	–.19	–2.08**
Environmental munificence	–.21	–.12	–1.35*
Adjusted R^2 = .10, F-statistic = 3.17** (d.f. = 6, 113)			

*Significant at $p < .10$ (one-tailed).
**Significant at $p < .05$ (one-tailed).
***Significant at $p < .01$ (one-tailed).

to new product market performance (H_{10}). This prediction is supported (b = –.04, t = –1.70, $p < .05$).

We also wanted to examine whether project inflexibility has any direct effect on performance. We regressed performance on project inflexibility along with an interaction term consisting of inflexibility and novelty. However, we did not find any effect.

Covariates for New Product Performance

As we expected, duration of the project increases market performance of the new product (b = .01, t = 3.10, $p < .01$). However, projects that spend a lot more money than other similar projects of the company are less successful (b = –1.04, t = –2.08, $p < .05$). Environmental munificence has a negative influence on market performance of the new product, though the effect is marginal (b = –.21, t = –1.35, $p < .1$).

Discussion and Conclusions

In this study, we examined some unintended consequences of Stage-Gate controls and showed that these controls have the potential to compromise some important aspects of the new product development process and to adversely affect novel new products. Our results suggest that when gate review criteria are more strict, objective (i.e., not open to interpretation or allowed to vary when the project requires), and frequently applied, they increase the inflexibility of the new project. The attempt to mitigate the adverse effect of rigorous criteria on project inflexibility through the adoption of conditional gates does not work. In turn, project inflexibility increases learning failure in the product development team. This adverse effect of project inflexibility on learning is worsened when there is turbulence in the technological sector of the environment. Learning failure due to Stage-Gate evaluation adversely affects the market performance of novel new products.

Among the hypotheses that have not been confirmed, we expected that gate conditionality would mitigate the adverse effects of gate review criteria on project inflexibility. We hypothesized this mitigating effect on the basis of the extant Stage-Gate and control systems literature, but the results do not confirm it. A possible explanation for the absence of this mitigating effect could be that Stage-Gate evaluation stabilizes various new product development routines, enhances the organization's faith in adherence to a prespecified sequence of product development activities, and promotes strict adherence to handover points between various highly interconnected activities (Benner and Tushman 2003; Cooper 1994). As a result, in reality, it becomes difficult to postpone finalizing details when the project moves from one set of activities to another, even though gate conditionality permits such postponement. Because we do not have data to support this argument, further research needs to examine this moderating effect of gate conditionality in detail.

Another hypothesis that has not been confirmed is the moderating effect of market turbulence on the relationship between project inflexibility and learning failure. Recall that we expected both market turbulence and technological turbulence to worsen the effect of project inflexibility on learning failure. Although we find no moderating effect of market turbulence, the effect of technological turbulence was confirmed. When project parameters become inflexible, perhaps it becomes more difficult to accommodate new technological information than market information. In general, new technological information is more likely to require basic changes in project parameters. Conversely, incorporating new market information may not always require a basic change in project parameters. As such, when new market information repeatedly emerges in the turbulent market environment, the project team does not get as frustrated, and in turn, its motivation to learn is not so adversely affected. Thus, learning failure may occur less often in the case of market turbulence. This is an ex post facto explanation for an unconfirmed effect and thus needs to be verified in further research.

Limitations

The following limitations of the study need to be kept in view while interpreting the findings: First, we gathered the perceptual measures used in this study from the same informant, which can introduce the possibility of common method bias. To detect the presence of this bias, Podsakoff and Organ (1986) recommend examining the items used to measure the different constructs for semantic overlap. In our study, there does not seem to be an obvious overlap in the items in different measures. Furthermore, to detect the presence of common method bias, we also conducted Harman's one-factor test (Podsakoff et al. 2003). The results of principal component analysis without rotation indicated that variables do not form a single higher-order factor, which suggests that common method bias is not a serious cause for concern. Then, the presence of moderator effects in our study reduces the possibility of common method bias (Podsakoff and Organ 1986). This study also has an objective variable (the number of gates), which mitigates percept–percept inflation (Crampton and Wagner 1994). Thus, it seems that the common method bias is not a major concern.

Second, there was some concern about social desirability bias. If a social desirability bias is present, responses should be heavily skewed toward the positive end of the scale, and the variation in responses should be limited (Andrews and Smith 1996; Podsakoff and Organ 1986). However, from the descriptive properties (i.e., mean and distribution) of the scales (Table 1), social desirability bias is not a cause for concern in this study.

Finally, our study only focuses on market performance as the outcome of new product development. If we had included cycle time and efficiency, we might have found that rigorous gate evaluation has a positive effect on such outcomes. Future studies might want to examine how rigorous gate controls influence outcomes other than market performance.

Research Implications

Our study has several implications for academic researchers. By drawing on research in process management, control systems, and organizational learning, we have provided a conceptual foundation to the Stage-Gate literature. So far, the Stage-Gate literature lacks a conceptual base. As we highlight, the Stage-Gate process should not be merely viewed as an attempt to introduce some checkpoints at various stages of product development. Gate evaluation sets into motion many subtle changes in new product development routines. The dynamics and focus of the product development activities change, and product development becomes more constrained. If future studies focus and draw on these conceptual underpinnings to study gate evaluation, they will help enrich and improve the underlying theory for this increasingly popular evaluation method.

Furthermore, we contribute to the Stage-Gate literature by showing that rigorous gate evaluation harms novel new products. The extant literature largely focuses on the benefits of gate controls. The main reservation about the utility of the Stage-Gate process has been that it cannot be used for game-changing radical innovations because such innovations use a nontraditional product development process to which gate evaluation cannot be applied. To date, what has been overlooked is the potential adverse effect of such controls on learning and the performance of novel products that use the traditional product development process.

In a related manner, another contribution of the study is in examining the effectiveness of the Stage-Gate process under conditions of environmental turbulence. We show that when there is turbulence in the technological sector of the firm's environment, rigorous gate controls have a worse effect on learning in the project and, thus, on the performance of novel products. In effect, our study enriches the literature by identifying two boundary conditions (i.e., novel products and technological turbulence) in which such gate controls harm rather than help the outcome of new product development activity.

Our study also extends the control systems literature by showing a trade-off between control and learning. The control literature has focused on dysfunctional effects of controls on employee behaviors (e.g., concealing and manipulating data) but has not examined whether controls can also undercut the effectiveness of organizations by impairing their ability to learn. In turn, this learning inability can erode the long-term competitiveness of the organization. We studied this trade-off in the context of the Stage-Gate process at the project level. Researchers on organizational control systems might want to replicate this effect at the organizational level. Furthermore, an implication of our research is that some of the concepts of control systems may not directly apply to a complex process such as Stage-Gate evaluation. For example, although some leeway or relaxation might help mitigate the adverse effect of tight controls, we show that conditional relaxation does not work for Stage-Gate evaluation.

This work is one of the first attempts to highlight the weaknesses of the Stage-Gate process and to demonstrate that an overreliance on this process might have an adverse effect on novel products. However, there is still a lot of ground that further research needs to cover related to Stage-Gate evaluation. For example, researchers need to explore alternative evaluation processes for new product development that might allow for more flexibility, particularly for novel products. Prior work has highlighted various product development processes other than the Stage-Gate process, such as a spiral process (McConnell 1996), evolutionary prototyping (McConnell 1996), and design to schedule and/or budget (Unger 2003). Although these different development processes have existed for some time, their frequency of use has been overshadowed by the Stage-Gate process. Further research could examine how these other processes affect project flexibility or what features can be borrowed from them to redesign the Stage-Gate process.

Finally, we studied the effects of rigorous gate controls at the project level. We chose the project-level focus because we believed that the Stage-Gate process is primarily a project-level control mechanism, and the effects of such controls are likely to manifest more distinctly at the project level. However, an equally valid approach would be to study the effectiveness of gate controls on firm-level outcomes. Firm-level research might be able to generate insights that cannot be generated at the project level. For example, only a firm-level study can provide information about how rigorous gate controls affect the performance of the new products of the firm taken together.

Managerial Implications

Our study has important implications for managers. The study raises concerns by highlighting that the Stage-Gate process has the potential of harming novel new products. Managers eagerly desire novel products because such products improve the likelihood of market success and profitability. At the same time, managers tend to have strong faith in the Stage-Gate process because it provides them with a useful tool for effective control of new product development in their firms. However, as our study suggests, it is problematic to aim for both rigorous gate controls and successful novel products simultaneously. Although rigorous gate controls may work for incremental products, they tend to harm the market performance of novel new products. Firms are likely to be better off designing separate gate controls for novel products (see Krubasik 1988). Even the gate review committee for more novel products needs to be different because a different mind-set is required for evaluating the progress of novel products compared with that of incremental products.

Next, the study warns managers that rigorous gate criteria are not appropriate when there is turbulence in the technology sector of the environment in which newer technical information can suddenly emerge during the process of product development. Rigorous gate criteria make projects inflexible, and such inflexibility in a turbulent environment worsens learning, which is crucial for novel new products. Thus, firms that operate in technologically volatile environments need to have different gate review criteria than firms that operate in relatively stable environments. Basically, the evaluation process needs to account for the special requirements of a turbulent environment and grant necessary latitude to new product development projects.

Our study also demonstrates to managers that merely tinkering with the Stage-Gate process may not solve the problem. For example, the recent suggestion of making the Stage-Gate process more conditional for mitigating some of the unintended consequences of rigorous gate evaluation is not effective. The gate evaluation process promotes strict adherence to handover points between various new product development activities. As a result, in reality, it becomes difficult to postpone finalizing details and meeting the criteria at a subsequent stage, even though gate conditionality permits such postponement. It is a better idea to create separate controls when the product is novel or the technological environment is turbulent.

In addition, while designing separate gate criteria for novel products, managers must realize that different novel projects can have different requirements and idiosyncrasies. Thus, it may not always be possible to come up with a set of standard criteria for evaluating all types of novel projects. A mechanism for taking care of differences in novel projects might be to let the product development team suggest how the standard criteria should be modified to accommodate the special needs of the novel project. The gate review committee can then evaluate the team's request and, after necessary negotiations with the team, can come up with customized evaluation criteria for that specific project. This customization can be done before the project formally goes through the evaluation process so that the team knows what is expected of it and the review committee is clear how and when the project will be evaluated over the course of its development.

Finally, Stage-Gate controls tend to become particularly rigorous in some firms because of the way they are implemented. Design guidelines provided by the Stage-Gate literature tend to predispose gate evaluation toward more rigor. In their enthusiasm to exert firm control, managers can sometimes make the evaluation process overly rigorous (Bechtold 1999; Merchant 1985). However, managers need to remember that the development of products that make a real difference in the marketplace requires the generation and effective implementation of novel ideas. Such implementation is hindered when controls are too constraining. Thus, it is in firms' long-term interest to check the tendency to make controls overly stringent.

In conclusion, our study highlights certain problems that are associated with the rigorous application of gate criteria, especially when the product is novel or when there is turbulence in the technological environment. In identifying these two boundary conditions, we have enhanced the literature on the Stage-Gate process for new product development and have cautioned managers to consider the context carefully while applying the Stage-Gate process.

APPENDIX
Measures

Construct	Items[a]
Strictly enforced criteria[a]	1. The criteria for gate reviews of new product development projects are formally established in your firm. 2. The criteria for gate reviews are formally used in your firm. 3. The criteria for gate reviews are strictly adhered to.
Objective criteria[a]	1. The criteria for gate reviews are open to interpretation. (reverse coded) 2. The criteria for gate reviews vary depending on who does the reviewing. (reverse coded) 3. The criteria for gate reviews are open to negotiations. (reverse coded)
Project inflexibility[a]	1. After approval at gate reviews, the approved project parameters were to be adhered to in a fairly rigid manner. 2. After approval, it was easy to change project parameters. (reverse coded) 3. After approval, the approved project definition became the rigid standard against which the progress was measured at the next gate review. 4. After approval, important project parameters were generally fixed with little room for negotiation.[d]
Gate conditionality[a]	1. Gate reviews of projects allow a project to be approved even if all criteria are not met, conditional on those criteria being met in the future. 2. The review criteria recognize that some projects follow a different development sequence and allow such projects to proceed further even if they only meet the criteria partially. 3. The review criteria recognize that sometimes various development activities need to be approved out of sequence and allow for that to happen. 4. At every stage, all project criteria must be met before a project is allowed to proceed further. (reverse coded)[d]
Learning failure[a]	1. After approval, project members had a tendency to focus less on new information (e.g., market- and/or technology-related information) that could have helped improve the project. 2. After the project was approved at gate reviews, it was difficult to incorporate relevant new information that would impact its definition. 3. After approval, even if project members felt that some changes were important, they knew it would be quite difficult to incorporate them in the product. 4. After approval, the focus was primarily on seeing the project sail smoothly through future gates, rather than on developing new insights for project improvement.[d]

APPENDIX
Continued

Construct	Items[a]
Technological turbulence[a]	1. The technology in our industry is changing rapidly. 2. Technological changes provide big opportunities in our industry. 3. A large number of new product ideas have been made possible through technological breakthroughs in our industry. 4. Technological developments in our industry are rather minor. (reverse coded)
Market turbulence[a]	1. It is difficult to predict how customers' needs and requirements will evolve in our markets. 2. It is difficult to forecast competitive actions. 3. Generally, it is difficult to understand how the market will change. 4. There is a great deal of uncertainty in our markets.
Product novelty[b]	As compared with competing products, please rate your product idea on each set of adjectives: •Predictable–novel •Usual–unusual •Ordinary–unique •Commonplace–original •Dull–exciting[d] •Old hat–trendsetting[d]
Product market performance[c]	Compared with the planned goals, how would you rate your product's: •Sales •Market share •Profit

[a] 7 = "strongly agree," and 1 = "strongly disagree."
[b] Seven-point semantic differential scale.
[c] Seven-point scale, anchored by "far below expected" and "far above expected."
[d] Item was dropped during scale refinement.

REFERENCES

Aiken, Leona S. and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage Publications.

Amabile, Teresa M. (1983), *The Social Psychology of Creativity*. New York: Springer-Verlag.

Andrews, Jonlee and Daniel C. Smith (1996), "In Search of Marketing Imagination: Factors Affecting the Creativity of Marketing Programs for Mature Products," *Journal of Marketing Research*, 33 (May), 174–87.

Armstrong, J. Scott and Terry S. Overton (1977), "Estimating Nonresponse Bias in Mail Surveys," *Journal of Marketing Research*, 14 (August), 396–402.

Atuahene-Gima, Kwaku (2005), "Resolving the Capability–Rigidity Paradox in New Product Innovation," *Journal of Marketing*, 69 (October), 61–83

Bechtold, Richard (1999), *Essentials of Software Project Management*. Norwood: MA: Management Concepts.

Benner, Mary J. and Michael Tushman (2002), "Process Management and Technological Innovation: A Longitudinal Study of the Photography and Paint Industries," *Administrative Science Quarterly*, 47 (December), 676–706.

——— and ——— (2003), "Exploitation, Exploration, and Process Management: The Productivity Dilemma Revisited," *Academy of Management Review*, 28 (April), 238–56.

Brown, Shona L. and Kathleen M. Eisenhardt (1995), "Product Development: Past Research, Present Findings, and Future Directions," *Academy of Management Review*, 20 (April), 343–79.

Canas, Jose J., Inmaculada Fajardo, and Ladislao Salmeron (2006), "Cognitive Flexibility," in *International Encyclopedia of Ergonomics and Human Factors*, Waldemar Karwowski, ed. Boca Raton, FL: CRC Press, 297–300.

Carbonell, Pilar, Ana Isabel Rodriguez Escudero, and Jose Luis Munuera Aleman (2004), "Technology Newness and Impact of Go/No-Go Criteria on New Product Success," *Marketing Letters*, 15 (July–October), 81–97.

Christensen, Clayton M. (1998), *The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail*. Boston: Harvard Business School Press

Cooper, Robert G. (1994), "Third-Generation New Product Processes," *Journal of Product Innovation Management*, 11 (January), 3–14

——— (1998), *Product Leadership: Creating and Launching Superior New Products*. Reading, MA: Perseus Books.

——— (2001), *Winning at New Products: Accelerating the Process from Idea to Launch*. New York: Perseus Publishing.

———, Scott J. Edgett, and Elko J. Kleinschmidt (2002a), "Optimizing the Stage-Gate Process: What Best Practices Companies Are Doing—Part I," *Research-Technology Management*, 45 (September–October), 21–27.

———, ———, and ——— (2002b), "Optimizing the Stage-Gate Process: What Best Practices Companies Are Doing—Part II," *Research-Technology Management*, 45 (November–December), 21–27

Crampton, Suzanne M. and John A. Wagner III (1994), "Percept–Percept Inflation in Micro-Organizational Research: An Investigation of Prevalence and Effect," *Journal of Applied Psychology*, 79 (February), 67–76.

Crawford, Merle C. (1993), "The Hidden Costs of Accelerated Product Development," *Journal of Product Innovation Management*, 9 (September), 188–99.

——— and Anthony DiBenedetto (2002), *New Products Management*. Chicago: Richard D. Irwin.

Danneels, Erwin (2002), "The Dynamics of Product Innovation and Firm Competences," *Strategic Management Journal*, 23 (December), 1095–1121.

Dess, Gregory G. and Donald W. Beard (1984), "Dimensions of Organizational Task Environments," *Administrative Science Quarterly*, 29 (March), 52–73.

Diamantopoulos, Adamantios and Heidi M. Winklhofer (2001), "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research*, 38 (May), 269–77.

Diamond, Nina and Zafar Iqbal (2000), "Ghosts of Projects Past: Exploring the Transfer of New Product Development Knowledge Between Projects," in *Marketing Theory and Applications*, Vol. 12, John Workman and William Perreault, eds. Chicago: American Marketing Association, 196.

Dougherty, Deborah and Trudy Heller (1994), "Illegitimacy of Successful Product Innovation in Established Firms," *Organization Science*, 5 (May), 200–218.

Duncan, Robert and Andrew Weiss (1979), "Organizational Learning: Implications for Organizational Design," in *Research in Organizational Behavior*, Barry Staw, ed. Greenwich, CT: JAI Press, 75–123.

Fiol, Marlene C. and Marjorie A. Lyles (1985), "Organizational Learning," *Academy of Management Review*, 10 (October), 803–813.

Garvin, David A. (1995), "Leveraging Processes for Strategic Advantage," *Harvard Business Review*, 73 (September–October), 77–90.

Griffin, Abbie (1997), "PDMA Research on New Product Development Practices: Updating Trends and Benchmarking Best Practices," *Journal of Product Innovation Management*, 14 (November), 429–58.

Hart, Susan, Erik Jan Hultink, Nikolaos Tzokas, and Harry R. Commandeur (2003), "Industrial Companies' Evaluation Criteria in New Product Development Gates," *Journal of Product Innovation Management*, 20 (January), 22–36.

Hedberg, Bo (1981), "How Organizations Learn and Unlearn," in *Handbook of Organizational Design*, Vol. 1, W.H. Starbuck and P.C. Nystrom, eds. New York: Oxford University Press, 3–27.

Hurley, Robert F. and G. Tomas M. Hult (1998), "Innovation, Market Orientation, and Organizational Learning: An Integration and Empirical Examination," *Journal of Marketing*, 62 (July), 42–53.

Hutt, Michael D. and Thomas W. Speh (2004), *Business Marketing Management: A Strategic View of Industrial Organizational Markets*. Mason, OH: South-Western.

Iansiti, Marco and Alan MacCormack (1997), "Developing Products on Internet Time," *Harvard Business Review*, 75 (September–October), 109–117.

Im, Subin and John P. Workman Jr. (2004), "Market Orientation, Creativity, and New Product Performance in High-Technology Firms," *Journal of Marketing*, 68 (April), 114–32

Jaccard, James, Choi K. Wan, and Robert Turrisi (1990), "The Detection and Interpretation of Interaction Effects Between Continuous Variables in Multiple Regression," *Multivariate Behavioral Research*, 25 (October), 467–78.

Jackson, Phillip W. and Samuel Messick (1965), "The Person, the Product, and the Response: Conceptual Problems in the Assessment of Creativity," *Journal of Personality*, 33 (September), 309–329.

Jaworski, Bernard (1988), "Toward a Theory of Marketing Control: Environmental Context, Control Types, and Consequences," *Journal of Marketing*, 52 (July), 23–39.

——— and Ajay K. Kohli (1993), "Market Orientation: Antecedents and Consequences," *Journal of Marketing*, 57 (July), 53–70.

Kanter, Rosabeth Moss (1983), *Change Masters*. New York: Simon and Schuster.

Krishnan, V. and Shantanu Bhattacharya (2002), "Technology Selection and Commitment in New Product Development: The Role of Uncertainty and Design Flexibility," *Management Science*, 48 (March), 313–27.

Krubasik, Edward G. (1988), "Customize Your Product Development," *Harvard Business Review*, 58 (November–December), 4–9.

Leifer, Richard, Christopher M. McDermott, Gina Colarelli O'Connor, Lois Peters, Mark Rice, and Robert W. Veryzer (2000), *Radical Innovation*. Boston: Harvard Business School Press.

Leonard, Dorothy (1992), "Core Capabilities and Core Rigidities: A Paradox in Managing New Product Development," *Strategic Management Journal*, 13 (Summer Special Issue), 111–25.

Levinthal, David A. and James G. March (1993), "The Myopia of Learning," *Strategic Management Journal*, 14 (Winter), 95–112.

MacCormack, Alan, Roberto Verganti, and Marco Iansiti (2001), "Developing Products on 'Internet Time': The Anatomy of a Flexible Development Process," *Management Science*, 47 (January), 133–50.

Mankin, Eric (2004), "Is Your Product-Development Process Helping or Hindering Innovation?" *Strategy and Innovation*, 2 (November–December), 3–5.

March, James G. (1991), "Exploration and Exploitation in Organizational Learning," *Organization Science*, 2 (February), 1–13.

McConnell, Steve (1996), *Rapid Development: Taming Wild Software Schedules*. Redmond, WA: Microsoft Press.

McDermott, Christopher M. and Gina Colarelli O'Connor (2002), "Managing Radical Innovation: An Overview of Emergent Strategy Issues," *Journal of Product Innovation Management*, 19 (November), 424–38.

Menon, Anil, Sundar G. Bharadwaj, Phani Tej Adidam, and Steven W. Edison (1999), "Antecedents and Consequences of Marketing Strategy Making: A Model and a Test," *Journal of Marketing*, 63 (April), 18–40.

Merchant, Kenneth A. (1985), *Control in Business Organizations*. Boston: Pitman.

——— (1988), "Progressing Toward a Theory of Marketing Control: A Comment," *Journal of Marketing*, 52 (July), 40–44.

Miner, Anne, Paula Bassoff, and Christine Moorman (2001), "Organizational Improvization and Learning: A Field Study," *Administrative Science Quarterly*, 46 (June), 304–337.

Moorman, Christine and Anne S. Miner (1998), "The Convergence of Planning and Execution: Improvisation in New Product Development," *Journal of Marketing*, 62 (July), 1–20.

Neter, John, William Wasserman, and Michael H. Kutner (1990), *Applied Linear Statistical Models*. Boston: Richard D. Irwin.

Nonaka, Ikujiro and Hirotaka Takeuchi (1995), *The Knowledge-Generating Company*. New York: Oxford University Press.

Podsakoff, Philip M., Scott B. MacKenzie, Nathan P. Podsakoff, and Jeong-Yeon Lee (2003), "Common Method Bias in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, 88 (October), 879–903.

——— and Dennis W. Organ (1986), "Self-Reports in Organizational Research: Problems and Prospects," *Journal of Management*, 12 (Winter), 531–44.

Ramaswami, Sridhar N. (1996), "Marketing Controls and Dysfunctional Employee Behaviors: A Test of Traditional and Contingency Theory Postulates," *Journal of Marketing*, 60 (April), 105–120.

Robinson, William T. and Jeongwen Chiang (2002), "Product Development Strategies for Established Pioneers, Early Followers, and Late Entrants," *Strategic Management Journal*, 23 (September), 855–66.

Sethi, Rajesh (2000), "Superordinate Identity in Cross-Functional Product Development Teams: Its Effect on New Product Performance and Its Antecedents," *Journal of the Academy of Marketing Science*, 28 (Summer), 330–44.

———, Daniel C. Smith, and C. Whan Park (2001), "The Effect of Cross-Functional Product Development Teams on the Innovativeness of New Consumer Products," *Journal of Marketing Research*, 38 (February), 73–85.

Sitkin, Sim, Kathleen M. Sutcliffe, and Roger G. Schroeder (1994), "Distinguishing Control from Learning," *Academy of Management Review*, 19 (July), 537–64.

Slater, Stanley F. and John C. Narver (1995), "Market Orientation and Learning Organization," *Journal of Marketing*, 59 (July), 63–74.

Smith, Preston G. and Donald G. Reinertsen (1991), *Developing Products in Half the Time*. New York: Van Nostrand Reinhold.

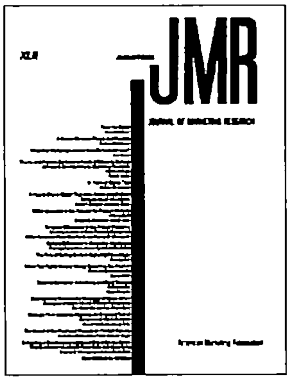
Unger, Darian W. (2003), "Product Development Process Design: Improving Development Response to Market, Technology, and Regulatory Risks," doctoral dissertation, Massachusetts Institute of Technology.

Verganti, Roberto (1999), "Planned Flexibility: Linking Anticipation and Reaction in Product Development Projects," *Journal of Product Innovation Management*, 16 (July), 363–76.

Veryzer, Robert W., Jr. (1998), "Discontinuous Innovation and the New Product Development Process," *Journal of Product Innovation Management*, 15 (July), 304–321.

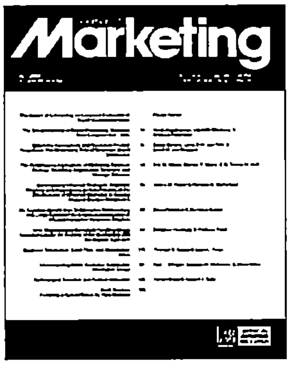
Zaltman, Gerald, Robert Duncan, and Jonny Holbek (1973), *Innovations and Organizations*. New York: John Wiley & Sons.

Check out our other AMA publications



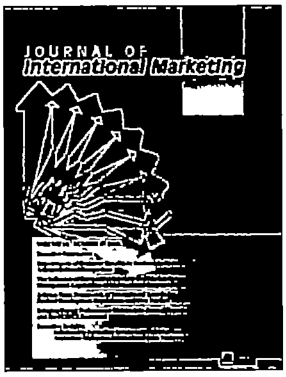
Journal of Marketing Research

Journal of Marketing Research is written for academics and practitioners of marketing research who need to be current with the latest marketing research techniques, methods, and applications of the marketing research function. *JMR* is a highly respected, quarterly publication that deals with the philosophical, conceptual, and technical aspects of the profession. Its readers include research analysts, educators, scholars, and statisticians.



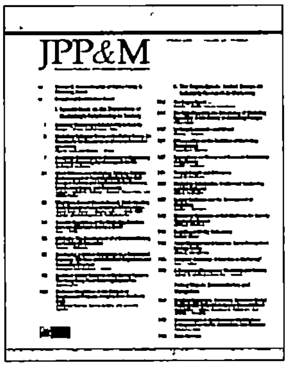
Journal of Marketing

Journal of Marketing was established in 1936, and is the premier publication for marketing academics and professionals in the field. It provides the marketing community with original research on all aspects of marketing. *Journal of Marketing* covers advertising, pricing, sales, distribution, ethics, and brand management, among other topics. This quarterly journal is designed to bridge the gap between marketing theory and practical application. Contributed articles are peer-reviewed by an experienced and highly respected board.



Journal of International Marketing

Journal of International Marketing is the foremost resource for understanding complex issues in today's international business environment. Written in a managerially relevant context, it provides important insights from the field by practicing international business executives. *Journal of International Marketing* also presents the most current research findings and studies conducted by top international scholars.



Journal of Public Policy & Marketing

Public policy has a profound influence on how products and services are marketed. In turn, marketing, by its very nature, influences public policy. *Journal of Public Policy & Marketing* addresses that dynamic, reciprocal relationship between marketing and public interest. It is read by public policymakers in federal, state, and local government agencies and by academics. *JPP&M* is peer reviewed for relevance and accuracy. Its review board represents not only many distinguished universities, but also the Federal Trade Commission and the Food and Drug Administration.

TO SUBSCRIBE call 800-262-1150 or go to www.amaorders.com

CALL NOW
to learn how AMA journals can keep you current with the most up-to-date marketing information.

NEW
MARKETING
TEXTBOOK

Don't make your students choose between:

- New textbook, price $150 or more
- Second hand textbook

*Assign them a **great textbook** at an **affordable** price*

$45 hard copy / **$25** electronic

Managing Marketing in the 21st Century

Examine ***Managing Marketing in the 21st Century*** online at **www.mm21c.com** (instructors' area available)

Noel Capon (author of *The Marketing Mavens*)
with
James Mac Hulbert
Columbia Business School

Wessex

www.mm21c.com

Journal of Marketing

March 2008
Volume 72, Number 2

Sheth Foundation/Journal of Marketing Award iii

Harold H. Maynard Award iv

Marketing Science Institute/H. Paul Root Award v

Customer Equity: An Integral Part of Financial Reporting 1 Thorsten Wiesel, Bernd Skiera, & Julián Villanueva

Demand-Based Pricing Versus Past-Price Dependence: A Cost–Benefit Analysis 15 Shuba Srinivasan, Koen Pauwels, & Vincent Nijs

Do Frontline Mechanisms Matter? Impact of Quality and Productivity Orientations on Unit Revenue, Efficiency, and Customer Satisfaction 28 Detelina Marinova, Jun Ye, & Jagdip Singh

Convergent Products: What Functionalities Add More Value to the Base? 46 Tripat Gill

Supply Chain Contagion 63 Richard G. McFarland, James M. Bloodgood, & Janice M. Payan

Trust at Different Organizational Levels 80 Eric (Er) Fang, Robert W. Palmatier, Lisa K. Scheer, & Ning Li

Determinants of Perceived Web Site Interactivity 99 Ji Hee Song & George M. Zinkhan

Innovation's Effect on Firm Value and Risk: Insights from Consumer Packaged Goods 114 Alina B. Sorescu & Jelena Spanjol

Configurations of Marketing and Sales: A Taxonomy 133 Christian Homburg, Ove Jensen, & Harley Krohmer

JOURNALS OF
THE AMERICAN MARKETING ASSOCIATION
AVAILABLE IN JSTOR

Since 2003, the American Marketing Association has worked with JSTOR, a not-for-profit online digital archive, to provide access to complete back runs of AMA journals and preserve them in a long-term archive. Users may currently search, browse, view, and print full-text PDF versions of articles in two AMA journals, from their first year published through all but the most recent five volumes:

Journal of Marketing (1936)
Journal of Marketing Research (1964)

These journals are available as part of JSTOR's Arts & Sciences Complement and Business II Collection. Users at institutions that participate in any of these collections can access the publications directly at ‹www.jstor.org›. The AMA publications *Journal of Public Policy & Marketing* and *Journal of International Marketing* will also be available in the JSTOR archive shortly.

Individual articles from the back issues of AMA's publications are available for purchase at $24.00 per article through JSTOR's Publisher Sales Service, a collaborative initiative to facilitate access to important scholarly literature. This program is meant to enable independent scholars unaffiliated with a participating institution or publisher to access single articles from journals archived in JSTOR. For more information on this program, please visit ‹http://www.jstor.org/about/pss_title_list.html›.

As the industry leader in providing expertise and knowledge to marketing professionals, the American Marketing Association is proud to collaborate with JSTOR to ensure the preservation of and access to the important scholarly literature of our field.

JSTOR is a not-for-profit organization helping the scholarly community discover, use and build upon a wide range of intellectual content in a trusted digital archive. Our efforts provide for the long-term preservation and access of leading academic journals and scholarly literature from around the world. Our work is supported by libraries, scholarly societies, publishers, and foundations.

Information regarding JSTOR available at www.jstor.org
149 Fifth Avenue, New York, NY 10010

Journal of
Marketing

Vol. 72, No. 2

CENTRAL LIBRARY
1-7-08
UNIVERSITY OF CALCUTTA
ALIPORE CAMPUS

AMERICAN MARKETING ASSOCIATION

EDITOR
Roland T. Rust
Robert H. Smith School of Business
3451 Van Munching Hall
University of Maryland
College Park, MD 20742-1815

ASSISTANTS TO THE EDITOR
Chris Bartone
E-mail: jom@ama.org

Millicent Locke
University of Maryland

WEB EDITOR
Siva K. Balasubramanian
Southern Illinois University

BOOK REVIEWS
Ramarao Desiraju
University of Central Florida

VICE PRESIDENT—PUBLICATIONS
Russell S. Winer
New York University

CHIEF EXECUTIVE OFFICER
Dennis Dunlap

GROUP PUBLISHER
Jack Hollfelder

EDITORIAL STAFF
Phone: (312) 542-9000
Fax: (312) 542-9001

MANAGING EDITOR
Francesca Van Gorp Cooley

TECHNICAL EDITOR
Andy Seagram

GRAPHICS EDITOR
Mary Loye

Advertising Staff
Fax: (312) 922-3763

Richard Ballschmiede, Advertising
(312) 542-9076

Sally Schmitz, Production
(312) 542-9038

JM Web site:
Submissions: www.marketingjournals.org/JM
Main page. www.marketingpower.com/JM

The *Journal of Marketing* (ISSN 0022-2429) is published bimonthly (January, March, May, July, September, and November). *Annual Subscription Rates*: Subscriptions are $115 for individuals in the United States and Canada and $160 in other countries. Library, school, and business subscriptions are $325 in the United States and Canada and $360 in other countries. Subscriptions for AMA members are $60 Prices are for one year. Single copy rate: $50 individual, $60 corporate in the United States and Canada; add $5 in other countries. Prepayment is required in U.S. dollars or the equivalent

Published by the Publications Group of the American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266 Periodicals postage paid at Chicago, IL and additional mailing offices. POSTMASTER· Send address changes to the *Journal of Marketing*, 311 S Wacker Dr., Suite 5800, Chicago, IL 60606-2266, USA.

•**MANUSCRIPTS**: Electronic submissions must be made via the *JM* Web site at http://www.marketingjournals.org. Paper submissions are not accepted. Editorial policies, review board members, Tables of Contents for *JM*, and so on, are available on the *JM* Web site at http://www marketingpower.com/jm.

•**ADVERTISERS AND ADVERTISING AGENCIES** assume liability for all content (including text, representations, and illustrations) of advertisements published, and also assume responsibility for any claims arising therefrom made against the publisher. The right is reserved to reject any advertisement.

•**SUBSCRIPTIONS***:* Members: Send to American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266 (1-800-AMA-1150 or 312/542-9000), or e-mail to info@ama.org. Nonmembers: Send to American Marketing Association, P.O. Box 11806, Birmingham, AL 35202, 1-800-633-4931, 205-995-1567, Fax. 205-995-1588, e-mail: amasubs@ebsco.com. Canadian residents add 6% GST (GST registration # 12747852).

•**PERMISSIONS AND REPRINT INFORMATION:** Without written permission from the AMA, any copying or reprinting (except by authors reprinting their own work) is prohibited. Requests for permission to reprint—such as copying for general distribution, advertising, or promotional purpose, for creating new collective works, or for resale—should be submitted in writing to permissions@ama.org. Bulk reprints are available by contacting Tamara Groft, Sheridan Reprints, (800) 352-2210 Permission is granted by AMA for those registered with the Copyright Clearance Center (CCC), 222 Rosewood Dr., Danvers, MA 01923, to photocopy any article herein for a flat fee. Contact CCC with ISSN 0022-2429, volume, and first and last page numbers of each article copied.

The views expressed in *JM* are not necessarily those of the editorial staff or the publisher. **All articles in this journal, unless otherwise noted, have undergone a blind review process.**

Copyright © 2008, American Marketing Association. All rights reserved

Editorial Review Board

ROHINI AHLUWALIA
University of Minnesota

KUSUM L AILAWADI
Dartmouth College

JOSEPH ALBA
University of Florida

TIM AMBLER
London Business School

EUGENE W ANDERSON
University of Michigan

TOR WALLIN ANDREASSEN
Norwegian School of Management

ASIM ANSARI
Columbia University

ERIC J. ARNOULD
University of Wyoming

KWAKU ATUAHENE-GIMA
CEIBS

BARRY L. BAYUS
University of North Carolina

WILLIAM BEARDEN
University of South Carolina

PAUL D BERGER
Boston University

PAUL N BLOOM
Duke University

LISA E. BOLTON
University of Pennsylvania

RUTH N. BOLTON
Arizona State University

ROBERT BORDLEY
General Motors Corporation

DOUGLAS BOWMAN
Emory University

RODERICK J BRODIE
University of Auckland

STEVEN P BROWN
University of Houston

JOSEPH CANNON
Colorado State University

DIPANKAR CHAKRAVARTI
University of Colorado

PIERRE CHANDON
INSEAD

YUXIN CHEN
New York University

ALEXANDER CHERNEV
Northwestern University

PRADEEP K. CHINTAGUNTA
University of Chicago

KEVIN CLANCY
Copernicus Marketing Consulting

PETER J DANAHER
University of Melbourne

KO DE RUYTER
Maastricht University

JOHN A DEIGHTON
Harvard University

MARNIK G. DEKIMPE
Tilburg University and Catholic
University Leuven

WAYNE S. DESARBO
Pennsylvania State University

RAVI DHAR
Yale University

NAVEEN DONTHU
Georgia State University

TÜLIN ERDEM
New York University

PETER S FADER
University of Pennsylvania

GARY L FRAZIER
University of Southern California

GARY F. GEBHARDT
University of South Florida

INGE GEYSKENS
Tilburg University

PETER N GOLDER
New York University

RAJDEEP GREWAL
Pennsylvania State University

ABBIE GRIFFIN
University of Utah

SUNIL GUPTA
Harvard University

ZEYNEP GURHAN-CANLI
KOC University

MICHAEL HAENLEIN
ESCP-EAP

JOHN R. HAUSER
Massachusetts Institute of
Technology

JAN B. HEIDE
University of Wisconsin–Madison

CHRISTIAN HOMBURG
University of Mannheim

WAYNE D. HOYER
University of Texas at Austin

JOEL HUBER
Duke University

JOHN HULLAND
University of Pittsburgh

G. TOMAS M HULT
Michigan State University

J JEFFREY INMAN
University of Pittsburgh

GANESH IYER
University of California, Berkeley

SANJAY JAIN
Texas A&M University

CHRIS JANISZEWSKI
University of Florida

SANDY D. JAP
Emory University

BERNARD J. JAWORSKI
Monitor Group

P K KANNAN
University of Maryland

TIMOTHY L KEININGHAM
Ipsos Loyalty

AMNA KIRMANI
University of Maryland

AJAY K. KOHLI
Emory University

ROBERT V KOZINETS
York University

V. KUMAR
University of Connecticut

DONALD R. LEHMANN
Columbia University

KATHERINE N LEMON
Boston College

MARY FRANCES LUCE
Duke University

XUEMING LUO
University of Texas at Arlington

ROBERT F. LUSCH
University of Arizona

JOHN G LYNCH JR.
Duke University

VIJAY MAHAJAN
University of Texas at Austin

RICHARD G MCFARLAND
Kansas State University

CARL F MELA
Duke University

VIKAS MITTAL
University of Pittsburgh

WENDY MOE
University of Maryland

CHRISTINE MOORMAN
Duke University

JOHN MOWEN
Oklahoma State University

EITAN MULLER
New York University and
Tel-Aviv University

SCOTT A NESLIN
Dartmouth College

RICHARD NETEMEYER
University of Virginia

RAKESH NIRAJ
University of Southern California

STEPHEN M. NOWLIS
Arizona State University

JOSEPH C. NUNES
University of Southern California

A. PARASURAMAN
University of Miami

KOEN PAUWELS
Dartmouth College

WILLIAM D PERREAULT
University of North Carolina

ROBERT PETERSON
University of Texas at Austin

F G M PIETERS
Tilburg University

JAIDEEP PRABHU
Imperial College

WILLIAM P. PUTSIS
University of North Carolina

ARVIND RANGASWAMY
Pennsylvania State University

AKSHAY RAO
University of Minnesota

VITHALA RAO
Cornell University

BRIAN T RATCHFORD
University of Texas at Dallas

WERNER REINARTZ
INSEAD

ARIC RINDFLEISCH
University of Wisconsin–Madison

FRED SELNES
Norwegian School of Management

STEVEN M. SHUGAN
University of Florida

ITAMAR SIMONSON
Stanford University

STANLEY F SLATER
Colorado State Universit

REBECCA SLOTEGRAAF
Indiana University

JOYDEEP SRIVASTAVA
University of Maryland

JAN-BENEDICT E M. STEENKAMP
University of North Carolina

DAVID W STEWART
University of Southern California

STEFAN STREMERSCH
Erasmus University Rotterdam

BAOHUNG SUN
Carnegie Mellon University

GERARD J TELLIS
University of Southern California

JACQUELYN THOMAS
Northwestern University

CRAIG THOMPSON
University of Wisconsin–Madison

DEBORA THOMPSON
Georgetown University

CHRISTOPHE VAN DEN BULTE
University of Pennsylvania

STIJN VAN OSSELAER
Erasmus University Rotterdam

P. RAJAN VARADARAJAN
Texas A&M University

STEPHEN L VARGO
University of Hawaii, Manoa

SAJEEV VARKI
University of South Florida

RAJKUMAR VENKATESAN
University of Virginia

PETER VERHOEF
University of Groningen

KENNETH WATHNE
University of Wisconsin–Madison

MICHEL WEDEL
University of Maryland

BARTON A WEITZ
University of Florida

RUSSELL S WINER
New York University

CAROLYN YOON
University of Michigan

ANTHONY ZAHORIK
Burke Institute

VALARIE A. ZEITHAML
University of North Carolina,
Chapel Hill

JIE ZHANG
University of Maryland

Z. JOHN ZHANG
University of Pennsylvania

SHETH FOUNDATION/JOURNAL OF MARKETING AWARD

The recipient of the 2007 Sheth Foundation/Journal of Marketing Award, which is administered through the American Marketing Association Foundation, is **Richard L. Oliver** for his article "Whence Consumer Loyalty," which appeared in the Special Issue 1999 (Vol. 63) "Fundamental Issues and Directions for Marketing" in *Journal of Marketing* (*JM*).

Nominations for the award were solicited from members of the *JM* Editorial Review Board, and a selection committee of *JM* former editors (Robert F. Lusch [chair], Ruth N. Bolton, and Thomas C. Kinnear) made the selection. The award is given to honor articles that have made long-term contributions to the *discipline of marketing*. The award recognizes scholarship based on the benefits of time and hindsight and acknowledges contributions and outcomes made to marketing theory and practice. This year, the committee considered all articles published in *JM* between 1998 and 2002. The committee also weighs information obtained from letters of nomination and citation analyses.

The article represents one of the best examples of rigorous theoretical reasoning and logic directed at a marketplace anomaly. The article has important managerial implications for customer relationship management in both business-to-consumer and business-to-business environments. The central issue addressed in the article is that though loyal customers may be highly satisfied, satisfied customers do not often translate into loyalty. Oliver demonstrates that various other factors, including personal determinism and social bonding at both the institutional and the personal level, explain this phenomenon. He concludes that for certain product categories or simply because of consumer disinterest, loyalty cannot be reasonably achieved. Importantly, "Whence Consumer Loyalty" has had a major impact on marketing research, as represented by the additional theoretical and empirical research it has stimulated. That impact is highly likely to continue into the future as further research helps clarify the mechanisms that work between satisfaction and loyalty, and vice versa.

The Sheth Foundation/Journal of Marketing Award was established through the generosity of the Sheth Foundation. The criteria for selection include the quality of the article's contribution to theory and practice, the article's originality, its technical competence, and its impact on the field of marketing. Articles eligible for the award must have been published at least five years but not more than ten years earlier.

The award was presented at the American Marketing Association's 2008 Winter Marketing Educators' Conference, held February 15–18, at the Hilton Austin in Austin, Texas.

•Richard L. Oliver is Professor of Management, Owen Graduate School of Management, Vanderbilt University.

HAROLD H. MAYNARD AWARD

Min Ding has been selected as the recipient of the 2007 Harold H. Maynard Award for his article "A Theory of Intraperson Games," which appeared in the April 2007 (Vol. 71, No. 2) issue of *Journal of Marketing*.

Nominations for the awards were solicited from members of the *JM* Editorial Review Board, and Jakki Thomas served as chair. The article was chosen for its significant contribution to *marketing theory and thought.*

The four (other) finalists were as follows: William T. Ross Jr. and Diana C. Robertson, for their article "Compound Relationships Between Firms" (July 2007); Kapil R. Tuli, Ajay K. Kohli, and Sundar G. Bharadwaj, for their article "Rethinking Customer Solutions: From Product Bundles to Relational Processes" (July 2007); Kent Grayson, for his article "Friendship Versus Business in Marketing Relationships" (October 2007); and Nermin Eyuboglu and Andreas Buja, for their article "Quasi-Darwinian Selection in Marketing Relationships" (October 2007).

The award was presented at the American Marketing Association's 2008 Winter Marketing Educators' Conference, held February 15–18, at the Hilton Austin in Austin, Texas.

•Min Ding is Associate Professor of Marketing, Smeal College of Business, Pennsylvania State University.

MARKETING SCIENCE INSTITUTE/H. PAUL ROOT AWARD

Timothy L. Keiningham, **Bruce Cooil**, **Tor Wallin Andreassen**, and **Lerzan Aksoy** have been selected as the recipients of the 2007 MSI/H. Paul Root Award for their article "A Longitudinal Examination of Net Promoter and Firm Revenue Growth," which appeared in the July 2007 (Vol. 71, No. 3) issue of *Journal of Marketing*.

Nominations for the awards were solicited from members of the *JM* Editorial Review Board, and Jakki Thomas served as chair. The article was chosen for its significant contribution to the *advancement of the practice of marketing*.

The three (other) finalists were as follows: Lien Lamey, Barbara Deleersnyder, Marnik G. Dekimpe, and Jan-Benedict E.M. Steenkamp, for their article "How Business Cycles Contribute to Private-Label Success: Evidence from the United States and Europe" (January 2007); Leigh McAlister, Raji Srinivasan, and MinChung Kim, for their article "Advertising, Research and Development, and Systematic Risk of the Firm" (January 2007); and Rex Yuxing Du, Wagner A. Kamakura, and Carl F. Mela, for their article "Size and Share of Customer Wallet" (April 2007).

The award was presented at the American Marketing Association's 2007 Winter Marketing Educators' Conference, held February 15–18, at the Hilton Austin in Austin, Texas.

- Timothy L. Keiningham is Senior Vice President and Head of Consulting, IPSOS Loyalty.
- Bruce Cooil is Professor of Management, Owen Graduate School of Management, Vanderbilt University.
- Tor Wallin Andreassen is Professor of Marketing and Chair, Department of Marketing, BI Norwegian School of Management.
- Lerzan Aksoy is Assistant Professor of Marketing, College of Administrative Sciences and Economics, Koç University.

Kohli Named as Next *JM* Editor

Ajay K. Kohli has been named editor of the *Journal of Marketing.* Kohli, whose three-year term begins July 1, 2008, replaces Roland T. Rust, who has served as editor since July 2005. Kohli is the Isaac Stiles Hopkins Chair in Marketing in the Goizueta Business School at Emory University. Kohli has taught at the undergraduate, graduate, and executive levels and has been recognized several times for teaching excellence. He is a recipient of the Schoolwide Jack G. Taylor Teaching Excellence Award at The University of Texas at Austin. He has led executive education seminars on topics such as building market-oriented organizations, differentiating commodities, and competing on customer value for firms in the United States, Europe, Asia, and Latin America.

Kohli's areas of specialization include market orientation, sales management, and business-to-business marketing. A major stream of Kohli's research focuses on what it means for a firm to be market oriented, how market orientation can be measured, why some firms are more market oriented than others, and what firms can do to become more market oriented. He has written extensively on the subject in journals such as *Journal of Marketing* and *Journal of Marketing Research.* His work subsequently has been reprinted in *Marketing Classics* and other books. His research on market orientation with Bernard Jaworski received the Alpha Kappa Psi award for the best practice-oriented article published in *Journal of Marketing* (1990), and his 1993 article in *Journal of Marketing* (also with Jaworski) received the inaugural Sheth Foundation/Journal of Marketing Award in 2002 for its long-term impact on the field of marketing.

Kohli is a member of the American Marketing Association, INFORMS, Academy of Marketing Science, and Association for Consumer Research, among others, and serves on the editorial boards of several journals, including *Journal of Marketing, Journal of Business Research,* and *International Journal of Research in Marketing.* He has consulted with many leading companies worldwide, including 3M, Accenture, Coca-Cola, Eastman Kodak, IBM, and Texas Instruments.

Thorsten Wiesel, Bernd Skiera, & Julián Villanueva

Customer Equity: An Integral Part of Financial Reporting

Recent initiatives demand information that supplements and complements a firm's financial statements to bridge the gap between financial statement capabilities and financial reporting objectives. Such information assists investors' decision making by explaining the main trends and factors that underlie the development, performance, and position of the firm's business. Firms that aim to increase the value of their customer base should report forward-looking customer metrics because such reports align customer management with corporate goals and investors' perspectives. The authors propose a means to report customer equity that enables investors—the "consumers" of financial reports—to monitor firms' performance with respect to their customer assets. Furthermore, they develop a specific model for Netflix and apply it to quarterly data from September 2001 to September 2006.

Keywords: customer equity, decomposition, financial reporting, customer lifetime value

The objective of financial reporting is to provide information to help current and potential investors, creditors, and other users (hereinafter, investors) assess the amounts, timing, and uncertainty of prospective cash receipts (Financial Accounting Standards Board [FASB] 1978; International Accounting Standards Board [IASB] 2004). However, the IASB's (2004) *Framework for the Preparation and Presentation of Financial Statements* acknowledges that financial statements (e.g., balance sheets, profit and loss statements, notes) are not, on their own, sufficient to meet the objective of financial reporting. To bridge the gap between what financial statements are able to achieve and the objective of financial reporting, firms must report additional information that explains the main trends and factors that underlie their development, performance, and position (IASB 2005).

In response, the "Management Discussion and Analysis" ("MD&A") required in the United States (Securities and Exchange Commission [SEC] 2003) and the recently discussed "Management Commentary" (IASB 2005) require information that supplements and complements information in a firm's financial statements. A recent report on the future of financial reporting published by the Big Six auditing firms confirms the importance of this discussion (Deloitte 2006). In Figure 1, we depict IASB's (2005, p. 12) view of financial reporting.

Specifically, the information requested in the "Management Commentary" should be future oriented, understandable, relevant, reliable, and comparable and should provide an "analysis through the eyes of management" (IASB 2005, p. 20). Examples of such information include details about the nature of the business, key resources, risks and relationships, and performance measures and indicators. Moreover, the IASB (2005) discussion paper explicitly mentions customer measures as crucial for assessing operating performance and, therefore, key information that should be reported to investors.

This call for more information is especially relevant for firms whose customers represent their primary assets because such firms aim to increase the value of their customer base through their customer management activities (e.g., Payne and Frow 2005; Reinartz and Kumar 2000, 2003; Reinartz, Thomas, and Kumar 2005; Rust, Lemon, and Zeithaml 2004; Ryals 2005; Venkatesan and Kumar 2004). Information about such activities exists within firms and is reviewed by management (e.g., Ambler 2000); if such information is important for managing the business, it also must be important to investors who want to assess performance and future prospects (PricewaterhouseCoopers 2005).

Consequently, firms should report forward-looking customer metrics (e.g., value of the customer base and its changes over time) in either the "MD&A" or "Management Commentary" sections to help investors—the "consumers" of financial reports—monitor firms' performance with respect to their customer assets and to communicate a customer value orientation to the financial community.[1] Such reporting would align with the described discussions about financial reporting and address recently mentioned research priorities focused on prescribing the critical marketing information elements that should be available to investors

Thorsten Wiesel is Assistant Professor of Marketing, Faculty of Economics and Business Administration, VU University Amsterdam (e-mail: wiesel@wiesel.info). Bernd Skiera is Professor of Marketing and a member of the board of the E-Finance Lab, School of Business and Economics, Johann Wolfgang Goethe–University Frankfurt/Main (e-mail: skiera@skiera.de). Julián Villanueva is Assistant Professor of Marketing, IESE Business School (e-mail: villanueva@iese.edu). This research was supported by the E-Finance Lab (Frankfurt), the Emory Marketing Institute, Meta Kaasch Stiftung, and Stiftung der Deutschen Wirtschaft. The authors gratefully acknowledge the three anonymous *JM* reviewers for their feedback and suggestions. The authors are also thankful for comments and suggestions provided by Sonja Gensler; Marc Fischer; Kees Camfferman; Eelke Wiersma; and seminar participants at the University of Groningen, University of Nijmegen, ESADE Business School, Emory University, and VU University Amsterdam.

---

[1] Note that we discuss an accounting measure of customer base value that can be used not in the primary financial statements (i.e., balance sheets) but rather in a supplementary section within financial reports (see Figure 1).

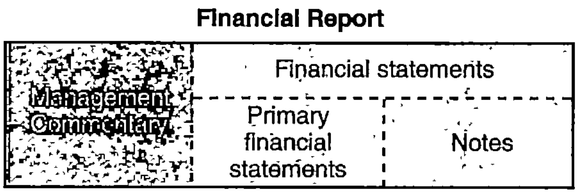
© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

(Srinivasan and Hanssens 2007). Furthermore, reporting may support marketing's reentry into the boardroom because it aligns customer management with corporate goals and the investor's perspective (McGovern et al. 2004). However, thus far, such information has not been reported, and only limited research has addressed how to report the value of the customer base and its changes over time.

External reporting about a firm's customer management activities must fall in line with financial reporting criteria and thus focus on the value of the customer base rather than concentrating on short-term-oriented value metrics, such as current profitability. Therefore, investors should receive information about (1) customer metrics (e.g., customer retention, customer cash flow), (2) the value of the customer base (usually operationalized as customer equity), (3) components of customer equity (e.g., customer equity before marketing expenditures, total lifetime retention expenditures, total lifetime acquisition expenditures), (4) changes in customer equity and components of customer equity over time, and (5) the effects of changes in customer metrics over time. Such data would provide valuable information to investors.

A simple and illustrative example demonstrates our motivation for reporting forward-looking customer metrics rather than just short-term metrics. Assume that a company with contractual relationships reports the metrics in the first four rows of Table 1 for two subsequent periods. The metrics indicate that management has done an excellent job, because they have significantly increased, which results in a boost in total cash flow by 31.43%. This kind of information is frequently reported. However, had the firm reported the next two rows of metrics (number of acquired customers and number of lost customers), overall assessments of this firm might change because the number of lost customers increased substantially, which leads to a much higher churn rate. If we consider the first eight rows of Table 1, evaluating whether management has done a good job is difficult because some metric changes are positive and others are negative. The overall effect remains unclear.

FIGURE 1
IASB's View of Financial Reporting



Using the available information to estimate a simple model of customer lifetime value (CLV) (for details, see Berger and Nasr 1998) shows that CLV diminished by 15.89%. Customer equity—defined here as CLV times the number of customers—also decreased by 7.87% (–$4,602.54). Moreover, decomposing the change in value (i.e., which sources are responsible for the 7.87% decline in customer equity; we discuss this decomposition subsequently) increases insights into and facilitates evaluations of management's performance. In the example, the increase in cash flow per customer leads to a positive change in customer equity of $11,690.28, but it is more than compensated for by a negative value effect due to the decreased retention rate (–$16,647.09).

This simple example illustrates that forward-looking customer metrics provide more—and, in this case, different—insights than short-term metrics. Instead of congratulating management for increasing the current period's cash flow by 31.43%, investors should ask management why it created short-term value at the expense of long-term value. Furthermore, the decomposition clarifies performance because it reveals the forward-looking consequences of changes in the short-term metrics and visualizes the sources of change in customer equity.

To extend from this example, our main objective in this research is to highlight the importance of reporting forward-looking customer metrics in a firm's financial report. This addresses the recent demand for additional information that

TABLE 1
Illustrative Example

	Period 1	Period 2	Change (%)
Cash flow per customer ($)	10.00	12.00	20.00
Total cash flow ($)	10,500	13,800	31.43
Total number of customers (beginning of period)	1000	1050	5.00
Total number of customers (ending of period)	1050	1150	9.52
Number of acquired customers (during the period)	150	300	100.00
Number of lost customers (during the period)	100	200	100.00
Churn rate (%)	9.76	18.19	86.37
Retention rate (%)	90.24	81.81	–9.34
CLV ($)	55.67	46.83	–15.89
Customer equity ($)	58,451.42	53,848.88	–7.87
Change in customer equity ($)		–4,602.54	
Due to changes in			
Cash flow per customer ($)		11,690.28	
Retention rate ($)		–16,647.09	
Number of customers ($)		5,566.80	
Other ($)		–5,212.53	

facilitates investors' decision making. In particular, we (1) emphasize the increasing need for forward-looking customer metrics to monitor customer management activities in financial reporting, (2) review a catalog of criteria relevant to financial reporting, (3) propose a technique to report the value of the customer base and its development over time, and (4) develop and apply a model that matches financial reporting criteria. Thus, we focus particularly on firms with contractual relationships (e.g., Internet service providers, financial service providers, telecommunication firms, energy suppliers, pay-television broadcasters, online movie rental services), which can easily determine the number of existing and lost customers at a particular point in time. Such determinations are more cumbersome for companies in noncontractual settings, which might need to modify this reporting technique.

We organize the remainder of this article as follows: The next section deduces a list of criteria relevant to financial reporting. Then, we propose a reporting technique that highlights the value of the customer base and its development over time. The subsequent section comprises a specific model that fits the proposed technique; we then apply this model to quarterly data from Netflix during 2001–2006. We conclude with implications and limitations of this model, as well as a discussion of further research.

Critical Criteria for Financial Reporting

The IASB's (2004) *Framework for the Preparation and Presentation of Financial Statements*, the *Statement of Financial Accounting Concepts No. 2* (FASB 1980), and the "Management Commentary" discussion paper (IASB 2005) all present various qualitative characteristics of financial reporting, as we list in the first column of Table 2. Because international accounting standards and the U.S. financial accounting standards are similar in many aspects, with boards currently working to converge them, we emphasize critical criteria common to both sets of standards.

TABLE 2
Financial Reporting Criteria

Qualitative Characteristics and Definitions	Derived Critical Criteria for Customer Equity Models
Relevance "The capacity to influence the economic decisions of users by helping them evaluate past, present, or future events or confirming or correcting their past evaluations."	Future orientation Decomposition
Reliability[a] "Information has the quality of reliability when it is free from material error, faithfully represents that which it either purports to represent or could reasonably be expected to represent, and is free from bias."	Objectivity
Comparability[b] "The quality of information that enables users to identify similarities in and differences between two sets of economic phenomena."	Comparability
Understandability "The quality of information that enables users to readily understand its significance."	Simplicity
Benefit > Cost	Cost effectiveness

[a] In the IASB discussion paper, reliability consists of supportability and balance. "Free from material error" and "represents faithfully" appear in supportability, and "free from bias" is part of balance.
[b] The IASB discussion paper concludes that the ability to compare management commentaries from an entity over time is important because comparability between entities as a qualitative characteristic conflicts with the objective of requiring management to convey what it believes is important.

Information is relevant if it influences decision making by the recipients of financial reports (e.g., analysts, investors, regulators) because it improves their predictions or verifies their prior expectations. Nagar and Rajan (2005) show empirically that a set of customer relationship measures improves the explanatory power for the subsequent year's earnings by 10%–15%, and Fornell and colleagues (2006) indicate that investments based on customer satisfaction produce sizable excess returns with lower systematic risk. Accounting journals reveal extant research that advocates nonfinancial measures of company performance, such as customer satisfaction and loyalty, as useful indicators of aspects of firm performance (e.g., Amir and Lev 1996; Said, HassabElnaby, and Wier 2003; Smith and Wright 2004). Ittner and Larcker (1998) find a significant, positive relationship between customer satisfaction measures and future performance and discover that announcing such information creates excess returns; in other words, the disclosure of customer satisfaction measures provides information to the stock market about expected future cash flow. Internally, nonfinancial measures provide managers with good indications of the reasons for change in, for example, customer equity, which again benefits the firm's performance. For example, Ittner, Larcker, and Randall (2003) reveal that financial service firms that use a broader set of financial and nonfinancial measures earn higher stock returns. According to these studies, as well as that of Gupta, Lehmann, and Stuart (2004), financial analysts have yet to pay more than scant attention to these off-balance-sheet assets, even though they may be key determinants of a firm's market value. If a firm were to include information about the health of its customer relationships, investors would gain a better understanding of the link between a firm's customer assets and the capacity to generate shareholder value. The IASB (2005) discussion paper explicitly mentions customer measures as significant indicators that, for example, retail banks might use to assess operating performance.

Any measure of the health of a firm's customer relationships should account for future cash flows, so the FASB

(2000) accentuates the importance of discounted cash flow techniques. Measures of CLV and customer equity meet this criterion because of their future orientation and similarity to discounted cash flow techniques. Decomposing changes in customer equity into its components and isolating the effects of changes in customer metrics enable investors to evaluate better the long-term impact of the current performance.

According to the IASB (2005) discussion paper, measures also must exhibit reliability, in that they are free from material error and bias and faithfully represent that which they either purport to represent or could reasonably be expected to represent. The "Management Commentary" discussion paper addresses the first and third elements as forms of supportability, such that a measure is supportable if it faithfully represents, for example, information sources, inherent uncertainty, and material assumptions, which enable investors to assess the reliability of the measure for themselves. A prerequisite is objectivity in data collection and processing, so that different people computing the measure will obtain the same value. Being free from bias means that the measures evenly assess good and bad aspects of the firm's performance and prospects.

Comparability represents another important quality of financial reporting information that implies that the measure applies consistently across industries and time, so investors can compare the results for different firms over several years. The "Management Commentary" discussion paper refers only to comparability over time because comparability between entities conflicts with the objective of requiring management to convey what it believes is important. However, a customer equity reporting technique could be applied and compared among entities if it were standardized and did not depend on specific types of data. Because this requirement also refers to the criterion of simplicity, we emphasize comparability both over time and across entities.

Furthermore, all measures should be readily understandable and cost effective. They should rely only on a few inputs and ideally use secondary information collected within the firm. Any necessary primary data collection should be reduced to a minimum because of the associated costs and lack of comparability.

In line with the preceding discussion, we emphasize six critical criteria: future orientation, decomposition, objectivity, comparability, simplicity, and cost effectiveness. Many of these criteria also appear in a list of desiderata developed during a Marketing Science Institute (1999) workshop intended for brand valuation purposes, a context in which Fischer (2007) similarly developed six brand asset valuation criteria for an accounting measure of brand equity. We followed his argumentation and extended it to be in line with our customer equity reporting purpose.

In summary, firms should report forward-looking customer metrics (i.e., value of the customer base and its changes over time) to help investors monitor firms' performance with respect to their customer assets. These forward-looking customer metrics must be consistent with the criteria presented in Table 2. Subsequently, we propose a technique to report the value of the customer base and its development over time that matches the financial reporting criteria.

Customer Equity Reporting

"Customer Equity Statement"

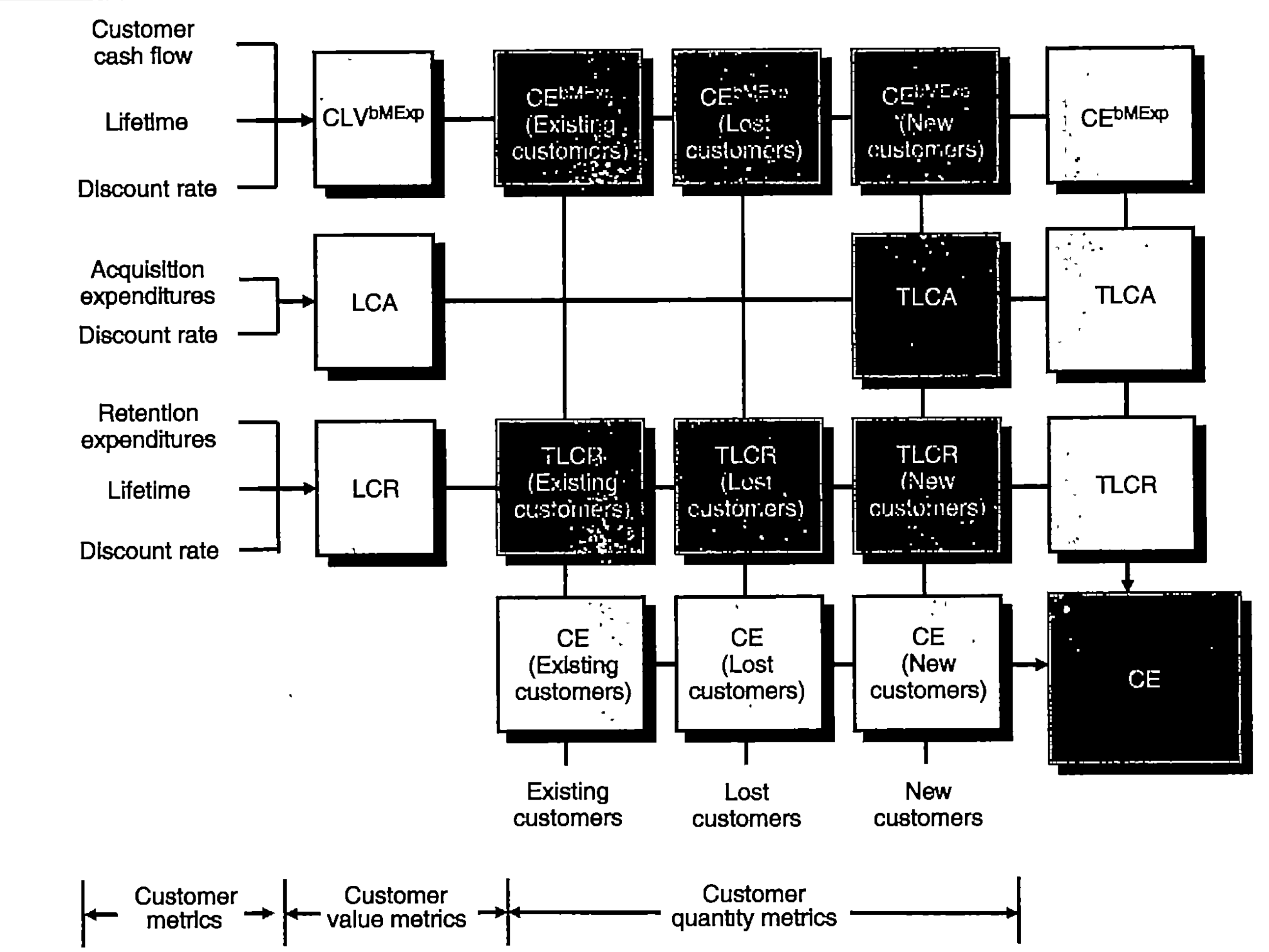
In general, customer equity reporting should comprise two main elements: the "Customer Equity Statement" and the "Customer Equity Flow Statement." The "Customer Equity Statement" reports customer equity (i.e., the value of the customer base) and its components in a single, clear display and thus reveals the value of the existing customer base. The "Customer Equity Flow Statement" describes changes in customer equity and its components between two periods and reports the influence of any changes in customer metrics on customer equity.

For the specific purpose of reporting, we define customer equity as the sum of the CLVs (after marketing expenditures) of all the firm's current customers in period t (Blattberg and Deighton 1996). The CLVs before marketing expenditures result from several customer metrics, such as cash flows generated by a customer (customer cash flow) and the duration of a customer's relationship with the company (customer lifetime). To retain or acquire customers, a firm must invest money; the measures of retention and acquisition expenditures per customer reflect those investments. Combining customer metrics with an appropriate discount rate provides a calculation of the net present value of a customer's cash flows (CLV before marketing expenditures), the net present value of a customer's acquisition expenditures (lifetime acquisition expenditures), and the net present value of a customer's retention expenditures (lifetime retention expenditures). We label these three metrics "customer value metrics" because they determine the value of a particular customer. Altogether, they determine each person's CLV (after marketing expenditures).

The number of customers at the end of a period equals the number of customers at the beginning of a period plus the number of customers acquired less the number of customers lost. To understand these customer movements, we use the number of existing customers (at the beginning of a period) and the number of new and lost customers (during a period) as "customer quantity metrics." Multiplying the CLV of an average customer before marketing expenditures by the number of existing, new, or lost customers provides the corresponding value of existing, new, or lost customers before marketing expenditures. We can perform a similar calculation for acquisition and retention expenditures. These various combinations of customer value and quantity metrics provide several different components of customer equity. As we illustrate in Figure 2, customer equity can be decomposed according to the kinds of customers (existing, new, or lost) or the value components (net present value of customer cash flows, retention expenditures, and acquisition expenditures).

In equation form, Figure 2 appears as follows, where the superscript E is existing customers at the beginning of the period, L is the number of lost customers during the period,

FIGURE 2
Decomposition of Customer Equity



Notes: CE = customer equity, CLV^bMExp = customer lifetime value before marketing expenditures, LCA = lifetime acquisition expenditures, LCR = lifetime retention expenditures, TLCA = total lifetime acquisition expenditures, and TLCR = total lifetime retention expenditures.

and N represents the number of new customers in the period:

$$(1)\quad CE_t = f_t^E(\cdot) - f_t^L(\cdot) + f_t^N(\cdot) - h_t^E(\cdot) + h_t^L(\cdot) - h_t^N(\cdot) - g_t^N(\cdot)$$
$$= CE_t^{bMExp} - TLCR_t - TLCA_t,$$

where

CE_t = customer equity (after marketing expenditures) at the end of period t,
$f_t^{E(L,N)}(\cdot)$ = customer equity before marketing expenditures of the existing (lost, new) customers at the end of period t,
$h_t^{E(L,N)}(\cdot)$ = net present value of retention expenditures for the existing (lost, new) customers during period t, and
$g_t^N(\cdot)$ = net present value of acquisition expenditures for the new customers during period t.

Thus, the sum of the first three terms in Equation 1 represents customer equity before marketing expenditures at the end of period t (CE_t^{bMExp}) and corresponds to customer equity before marketing expenditures at the beginning of the period (first term) less the customer equity of lost customers (second term) plus the customer equity of newly acquired customers (third term). The subsequent three terms represent the net present value of the retention expenditures ($TLCR_t$), again divided into corresponding values for customers with the company at the beginning of the period, those acquired, and those lost. The last term in Equation 1 represents the net present value of the acquisition expenditures ($TLCA_t$). Furthermore, we might rearrange Equation 1 to represent the value of different groups of customers, as follows:

$$(2)\quad CE_t = f_t^E(\cdot) - h_t^E(\cdot) - \left[f_t^L(\cdot) - h_t^L()\right] + f_t^N(\cdot) - h_t^N(\cdot) - g_t^N(\cdot)$$
$$= CE_t^E - CE_t^L + CE_t^N,$$

where

$CE_t^E = f_t^E(\cdot) - h_t^E(\cdot)$ = customer equity (after marketing expenditures) of existing customers at the end of period t,
$CE_t^L = f_t^L(\cdot) - h_t^L(\cdot)$ = customer equity (after marketing expenditures) of lost customers at the end of period t, and

$CE_t^N = f_t^N(\cdot) - h_t^N(\cdot) - g_t^N(\cdot)$ = customer equity (after marketing expenditures) of new customers at the end of period t.

Equation 2 states that customer equity (after marketing expenditures) at the end of the period equals the customer equity of existing customers after marketing expenditures (CE_t^E) less customer equity after marketing expenditures of customers lost during the period (CE_t^L) plus customer equity after marketing expenditures of newly acquired customers (CE_t^N). Thus, we decompose customer equity according to its value components (customer cash flows, retention expenditures, and acquisition expenditures) and groups of customers (existing, new, or lost). A wide range of customer equity models, such as those proposed by Blattberg and Deighton (1996), Rust, Lemon, and Zeithaml (2004), or Gupta, Lehmann, and Stuart (2004), can be used to specify $f_t(\cdot)$, $h_t(\cdot)$, or $g_t(\cdot)$; for a recent review of different customer equity models, see Kumar and George (2007).

"Customer Equity Flow Statement"

The "Customer Equity Flow Statement" illustrates changes in customer equity between two periods—that is, the difference between customer equity at the end of period t and t – 1:

$$(3) \qquad \Delta CE_{t,t-1} = CE_t - CE_{t-1}.$$

We can easily extend Equation 3 to calculate changes in customer equity among more than two periods. If we combine it with Equation 1, we further can decompose changes in customer equity into differences according to the number of existing, new, or lost customers. Moreover, knowledge of the exact specification of $f_t(\cdot)$, $h_t(\cdot)$, or $g_t(\cdot)$ would enable us to determine the effect of customer metrics on changes in customer equity.

Application

Objectives

We apply our reporting technique to Netflix. In doing so, we provide a possible specification of $f_t(\cdot)$, $h_t(\cdot)$, or $g_t(\cdot)$ that suits the available data. However, note that the "Customer Equity Statement" and the "Customer Equity Flow Statement" could handle other specifications, which we address in the "Discussion and Conclusion" section. Netflix's principal activity is to provide online movie rental services through access to more than 55,000 movies, television, and other entertainment titles. The standard subscription plan gives customers up to three titles at the same time with no due dates, late fees, or shipping charges. Shipping and receiving centers throughout the United States deliver the DVDs through the U.S. Postal Service at no charge to customers.

Because Netflix is listed on the NASDAQ, it must fulfill several SEC requirements, such as the "MD&A" section in its financial reports. In its 10-Q statements, Netflix (2006) provides information about customer churn and customer acquisition costs and notes that management not only reviews churn rates to evaluate whether the company is retaining existing customers, in accordance with its business plans, but also reviews acquisition expenditures to evaluate the efficiency of marketing programs for acquiring new customers. Furthermore, the statements indicate that Netflix believes in the usefulness of monitoring these metrics together, not individually, because it will not make business decisions based on a single metric. Although these metrics are measurable and observable by investors over time, they contradict several requirements of financial reporting; their interpretation is far from simple because investors must trade off between any changes in the metrics to determine the overall, long-term effect.

Data

We use publicly available, quarterly data from annual reports, 10-K and 10-Q statements, and other company reports from September 2001 to September 2006. The data for each quarter include the number of customers, average monthly churn rate, gross subscriber additions, subscription revenue, subscription costs of revenue, operating expenses (without marketing expenditures), acquisition cost per customer, and marketing expenditures.

This information enables us to calculate the number of lost customers during the quarter (i.e., difference in the number of customers between the current and previous quarter plus gross additions in the current quarter). We calculate the average quarterly retention rate as 1 – (number of lost customers during the quarter divided by the average number of customers during the quarter). To calculate the cash flow per customer, we subtract the subscription costs of revenue and the operating expenditures from the subscription revenue and divide the result by the number of customers. To smooth out seasonal fluctuations and other one-off effects, we use a common financial practice called a "trailing 12-month average" and replace the cash flow per customer and quarter, as well as the quarterly retention rate, with the mean of the corresponding cash flows and retention rates for the preceding four quarters. Multiplying the number of customers acquired by the customer acquisition cost yields the acquisition expenditures, which equal the marketing expenditures of Netflix; that is, Netflix considers all its marketing expenditures acquisition expenditures. Thus, Netflix claims that its retention expenditures are zero.[2] The company provides no information about its discount rate, so we choose an annual discount rate of 10% (quarterly discount rate amounts to 2.41%).

In Table 3, we list our calculations for the different customer metrics, as well as the resulting customer metrics for the preceding four quarters (Q4 2005–Q3 2006). Figure 3 illustrates the value and changes of customer metrics over time. On the positive side, Netflix increased its number of customers and its retention rate, as well as the cash flow per customer, in 2006 after suffering a drop in 2005. However, its acquisition expenditures increased. Thus, these measures

---

[2] Some of Netflix's technology and development expenses may result from the "member" portion of its Web site and thus reflect retention expenditures. Unfortunately, it is difficult to disentangle acquisition, retention, and general purpose of these expenditures. Therefore, we do not use them to determine retention expenditures.

TABLE 3
Calculation of Customer Metrics

	Calculation Method or Data Source	Q4 2005	Q1 2006	Q2 2006	Q3 2006
Number of customers (in thousands)	Reported number of customers per quarter (source: financial statements)	4179	4866	5169	5662
Number of new customers (in thousands)	Reported number of gross subscriber additions (source: financial statements)	1156	1377	1070	1310
Number of lost customers (in thousands)	Difference in number of customers between the current and the previous quarter + number of gross additions in the current quarter	569	690	767	817
Customer cash flow ($)	(Subscription revenue – subscription cost of revenue – operating expenses without marketing)/number of customers	9.97	10.84	11.87	12.60
Retention rate	1 – [number of lost customers during quarter/(number of customers at the beginning of quarter + number of customers at the end of quarter)/2]	.83	.84	.85	.85
Retention expenditures per customer ($)	(Reported marketing expenditures – reported acquisition cost per customer × number of new customers)/(number of customers – number of new customers)	.00	.00	.00	.00
Acquisition expenditures per customer ($)	Reported acquisition cost per customer (source: financial statements)	38.08	38.47	43.95	45.32

do not provide a clear picture of the overall value of the customer base. To obtain that picture, we must specify $f_t(\cdot)$, $h_t(\cdot)$, or $g_t(\cdot)$ in accordance with the available data, as we do in the next subsection.

Model

Although our formulation in Equations 1 and 2 is flexible enough to capture a wide range of specifications, data availability limits the feasibility of some models in our application. Our available information allows us to calculate only average values, such as average cash flow per customer, so we build on Berger and Nasr's (1998) ideas and select a parsimonious, easily applicable CLV specification (see also Gupta and Lehmann 2003; Kumar, Ramani, and Bohling 2004). This specification provides an example of a particular formulation of a customer equity model for financial reporting purposes; additional information, such as the metrics according to different customer segments, would enable us to capture heterogeneity across different customer segments and use alternative model formulations.

Calculating CLV before marketing expenditures, lifetime retention expenditures, and lifetime acquisition expenditures. The CLV before marketing expenditures of a current customer i in period t ($CLV_{i,t}^{bMExp}$) reflects the present value of the customer's cash flows ($C_{i,t+t'}$) over his or her remaining lifetime ($T_i - t$), with k as the discount rate:

$$(4)\qquad CLV_{i,t}^{bMExp} = \sum_{t'=0}^{T_i - t} \frac{C_{i,t+t'}}{(1+k)^{t'}}.$$

To develop an average CLV_t^{bMExp}, we assume that retention rates, which reflect the behavior of an average customer, are constant over time ($T_i \rightarrow r_t$) and that the customer cash flows of an average customer are also constant over time ($C_{i,t+t'} = C_t$). Using these assumptions, we can rewrite Equation 4 to describe the lifetime value of an average customer:

$$(5)\qquad CLV_t^{bMExp} = \sum_{t'=0}^{\infty} \frac{C_t \times r_t^{t'}}{(1+k)^{t'}} = \frac{C_t \times r_t^0}{(1+k)^0} + \frac{C_t \times r_t^1}{(1+k)^1} + \dots$$

Because Equation 4 is an infinite geometric series for $|r/(1 + k)| < 1$, we also can rewrite it as

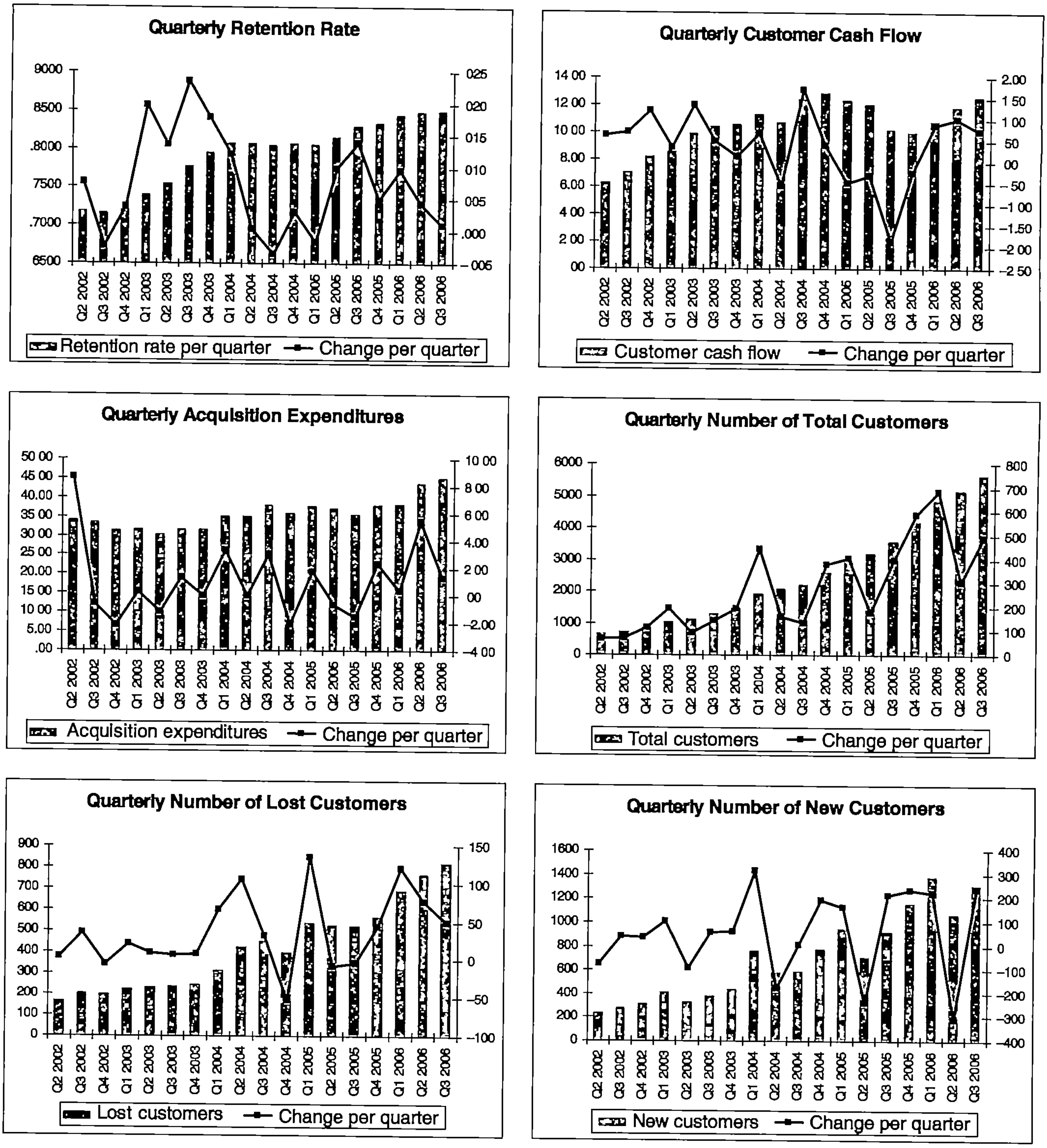
$$(6)\qquad CLV_t^{bMExp} = C_t \times \frac{1+k}{1+k-r_t}.$$

According to Equation 6, we recognize that lifetime retention expenditures (LCR_t) equal Equation 7 in the case of constant retention expenditures over time (CR_t). Thus, we assume that retention expenditures occur in all periods:

$$(7)\qquad LCR_t = CR_t \times \frac{1+k}{1+k-r_t}.$$

The lifetime acquisition expenditures (LCA_t) equal the current acquisition expenditures (CA_t) if they occur only in the acquisition period. Otherwise, we must calculate the net present value of the acquisition expenditures. The CLV after marketing expenditures for existing customers, $CLV_{E,t}^{bMExp}$,

FIGURE 3
Development of Customer Metrics Over Time



Quarterly Retention Rate

Quarterly Customer Cash Flow

Quarterly Acquisition Expenditures

Quarterly Number of Total Customers

Quarterly Number of Lost Customers

Quarterly Number of New Customers

is CLV_t^{bMExp} less LCR_t, whereas that for new customers, $CLV_{N,t}^{bMExp}$, equals CLV_t^{bMExp} less the sum of LCA_t and LCR_t.

To assess Netflix, this particular formulation for calculating CLV works well and fulfills the criteria for financial reporting. However, we note the potential issues involved in our use of constant retention rates and the "gone-for-good" assumption; both Kumar and Reinartz (2005) and Fader and Hardie (2006) show that using Equation 6 to calculate CLV might underestimate the value of the customer base if customer retention rates are heterogeneous. In that case, Equation 6 might include a factor that adjusts for the heterogeneity of retention rates across customers (Fader and Hardie 2006). Alternatively, we could estimate the equation separately for homogeneous customer segments. Neither approach is applicable to our illustration because of the limited information availability, but they could be adopted easily if additional data were available. Although this limitation might bias the results, we note that the main objective of the "Customer Equity Flow Statement" is to track the development of customer equity over time. Therefore, if the heterogeneity of retention rates remains constant over time, the bias largely cancels itself out in the relative comparisons across different points in time. In addition, strong heterogeneity underestimates both CLV before marketing expenditures (CLV_t^{bMExp}) and lifetime retention expenditures (LCR_t). Thus, the amount of bias should cancel out at least partly because CLV after marketing expenditures subtracts LCR_t from CLV_t^{bMExp}.

Rust, Lemon, and Zeithaml (2004) also demonstrate the gone-for-good assumption bias, in that customers may return. This assumption biases calculations in two ways: First, it may underestimate the value of a customer (before marketing expenditures), who might eventually return and generate new cash flows. Second, it can underestimate acquisition expenditures for new customers because some "new" customers may be "reactivated" customers, in which case it may be more expensive to convince them about service improvements, for example. Acquisition expenditures related to these customers also may be overestimated if customers cancel the service as a result of a temporary absence (i.e., being abroad for six months). Netflix does not disclose any further information about the number of new customers and acquisition expenditures, but we posit that these opposing effects on CLV (before marketing expenditures) and acquisition expenditures should cancel each other out at least partly and thus limit the possible bias of the gone-for-good assumption.

"Customer Equity Statement." When we combine Equations 1, 6, and 7, we obtain the following specification for customer equity (CE_t) at the end of period t:

$$\text{(8)} \qquad CE_t = N_t \times C_t \times \frac{1+k}{1+k-r_t} - N_t \times CR_t \times \frac{1+k}{1+k-r_t} - N_t^N \times CA_t.$$

The first term on the right-hand side of Equation 8 represents our specification of $f_t(\cdot)$, the second term represents our specification of $h_t(\cdot)$, and the third term represents our specification of $g_t(\cdot)$. We choose this specification because its suits the available data, but we note that our reporting technique is also able to capture other specifications, which we address in the "Discussion and Conclusion" section. We determine customer equity according to customer metrics (C_t, r_t, CR_t, and CA_t) and customer quantity metrics (N_t and N_t^N). Using the concepts we described previously, it is straightforward to build on the ideas of Equations 1 or 2 and to further decompose Equation 8 into customer equity components.

"Customer Equity Flow Statement." The "Customer Equity Flow Statement" provides information about absolute changes in customer equity and its components over time. Equation 9 determines the change in customer equity between periods t and t – 1 and can be easily modified to calculate changes in the components of customer equity:

$$\begin{aligned}\text{(9)} \quad \Delta CE_{t-1,t} &= CE_t - CE_{t-1} \\ &= \Bigg[\bigg(N_t \times C_t \times \frac{1+k}{1+k-r_t} - N_t \times CR_t \\ &\quad \times \frac{1+k}{1+k-r_t} - N_t^N \times CA_t\bigg) - \bigg(N_{t-1} \\ &\quad \times C_{t-1} \times \frac{1+k}{1+k-r_{t-1}} - N_{t-1} \times CR_{t-1} \\ &\quad \times \frac{1+k}{1+k-r_{t-1}} - N_{t-1}^N \times CA_{t-1}\bigg)\Bigg].\end{aligned}$$

The influence of any particular changes in customer metrics on customer equity appears in the difference in customer equity that results from a change in that particular metric. For example, the influence of changes in the number of existing customers between periods t and t – 1 on customer equity before marketing expenditures is as follows:

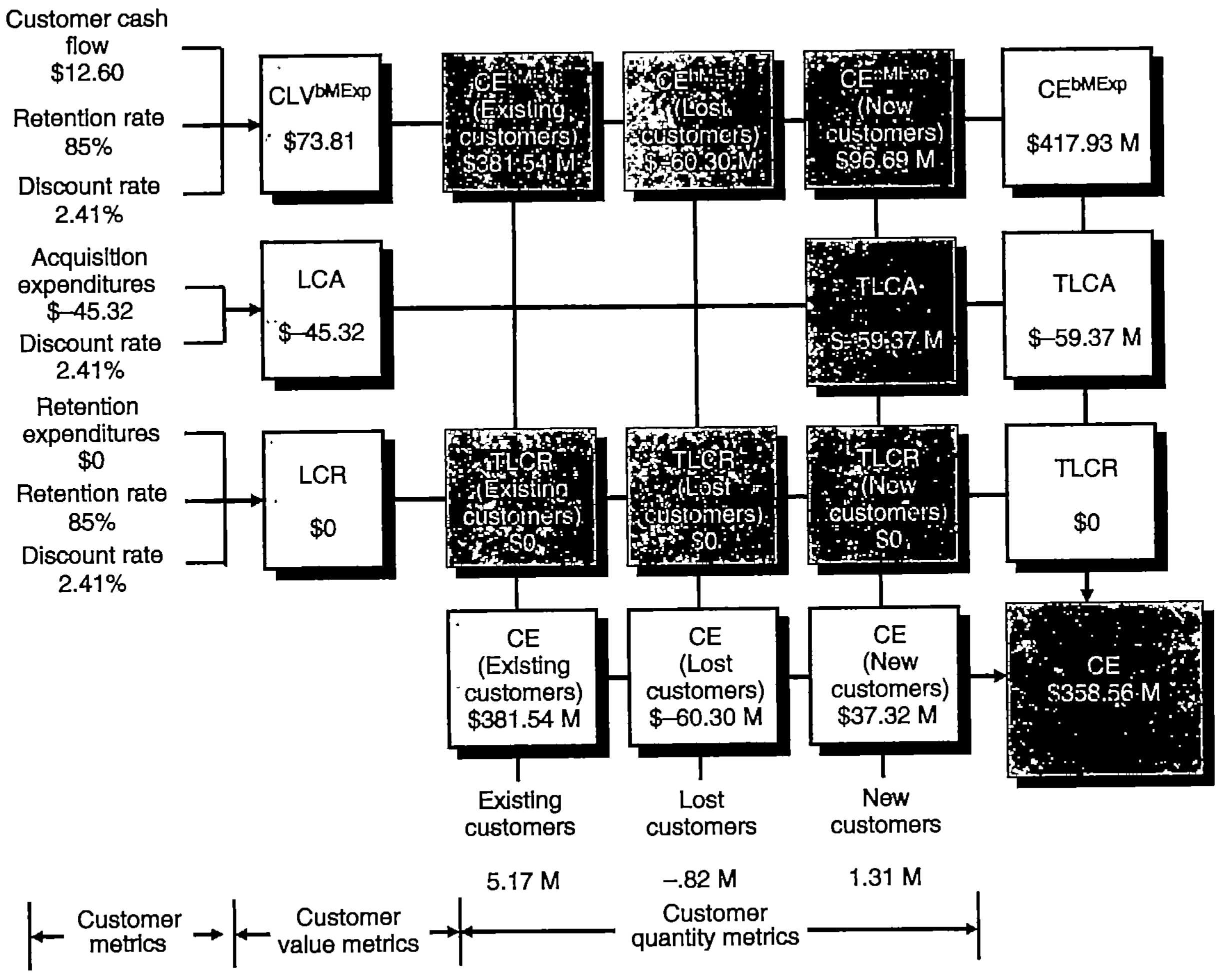
$$\begin{aligned}\text{(10)} \quad \Delta N_{t-1,t}^{E,f} &= f_t^E\left(N_{t-1}, C_{t-1}, L_{t-1}\right) \\ &\quad - f_{t-1}^E\left(N_{t-2}, C_{t-1}, L_{t-1}\right) \\ &= N_{t-1} \times C_{t-1} \times \frac{1+k}{1+k-r_{t-1}} \\ &\quad - N_{t-2} \times C_{t-1} \times \frac{1+k}{1+k-r_{t-1}} \\ &= \left(N_{t-1} - N_{t-2}\right) \times C_{t-1} \times \frac{1+k}{1+k-r_{t-1}}.\end{aligned}$$

The model for reporting customer equity is consistent with the criteria presented in Table 2: It accounts for future cash flows (future orientation), decomposes customer equity and changes in customer equity (decomposition), uses accounting or other objective data as inputs (objectivity), is standardized and does not depend on specific data (comparability), is easy to understand for decision makers (simplicity), and does not require costly data collection because it uses data already available in the firm (cost effectiveness).

Netflix's "Customer Equity Statement"

In Figure 4, we depict Netflix's "Customer Equity Statement" for Q3 2006, on the basis of Figure 2. Customer equity yields \$358.56 million in Q3 2006, according to the customer equity without marketing expenditures for existing customers (\$381.54 million), lost customers (–\$60.30 million), new customers (\$96.69 million), and total lifetime acquisition expenditures (–\$59.37 million). Because Netflix spends all its marketing expenditures on acquiring new customers, the total lifetime retention expenditures are always zero. We also demonstrate the decomposition according to groups of customers in Figure 4.

FIGURE 4
Netflix's Customer Equity Statement (Q3 2006)



Notes: CE = customer equity, CLV^{bMExp} = customer lifetime value before marketing expenditures, LCA = lifetime acquisition expenditures, LCR = lifetime retention expenditures, TLCA = total lifetime acquisition expenditures, and TLCR = total lifetime retention expenditures.

In Figure 5, we illustrate the development of customer equity over time. Except for Q3 2005, customer equity always increases over time. Because it monitors customer equity over time, the "Customer Equity Statement" provides investors with information about the value of the customer base, as well as an illustrative overview of customer metrics, the value of the customer base, and its components. However, it does not indicate the sources of change in customer equity over time, which would enhance any analysis by giving investors insights into how much and, as a result, of which metric the value of the customer base has changed. More detailed statements about the firm's customer management activities appear in the "Customer Equity Flow Statement."

Netflix's "Customer Equity Flow Statement"

Following from Figure 4, we develop Figure 6 to depict Netflix's total change in customer equity, its components, and its customer metrics in Q2–Q3 2006. Customer equity changed by $48.10 million, which reflects a change in customer equity before marketing expenditures of $60.44 million and a change in total lifetime acquisition expenditures of –$12.34 million (i.e., total lifetime acquisition expenditures increased). The change in customer equity before marketing expenditures comprises three components: change in

FIGURE 5
Customer Equity Over Time

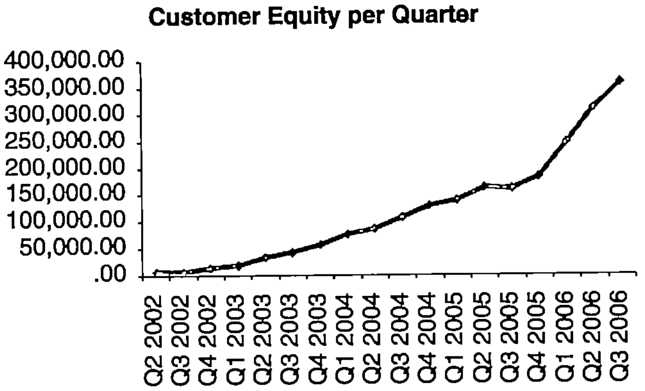
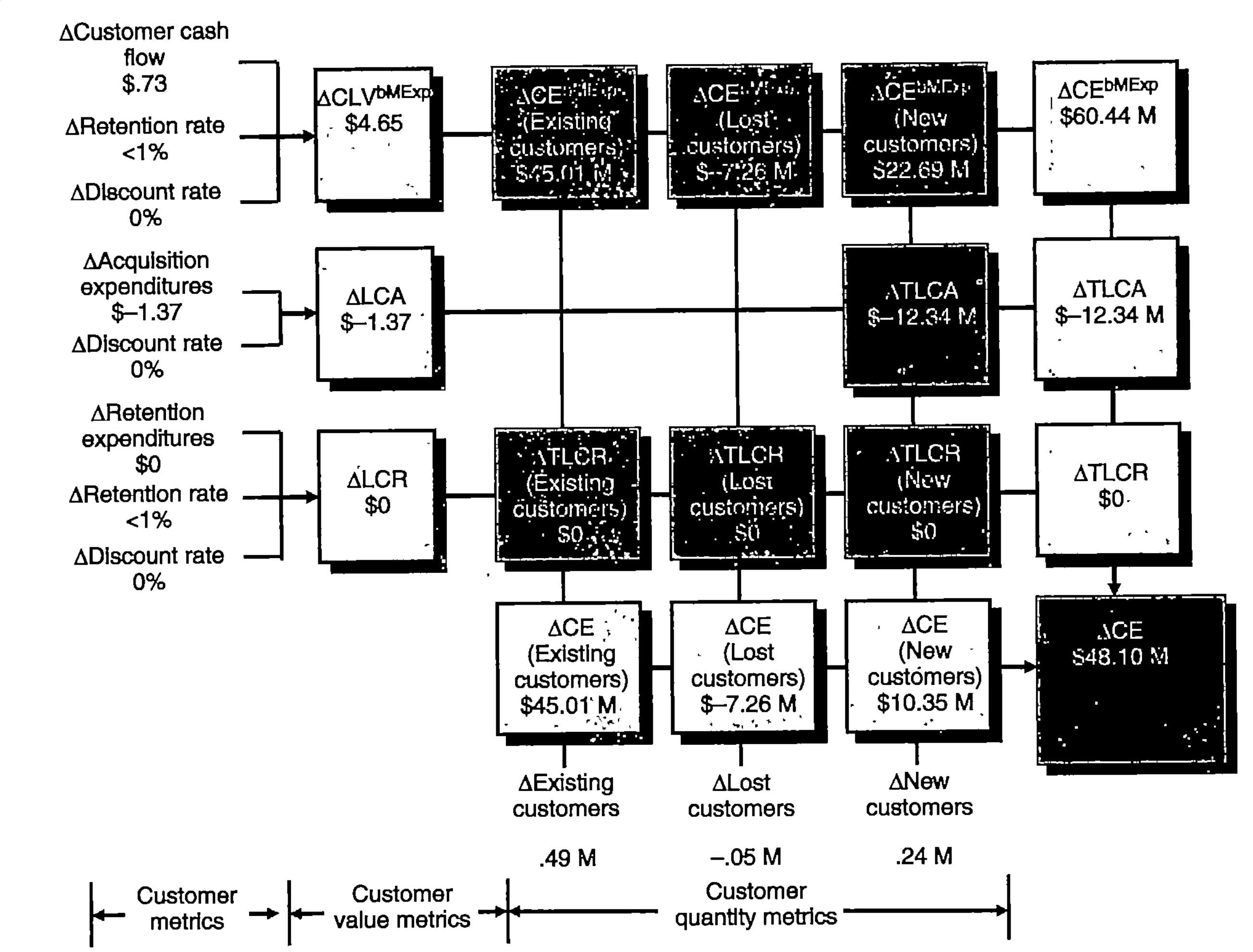


FIGURE 6
Netflix's Customer Flow Equity Statement (Q2–Q3 2006): Components View



Notes: ΔCE = customer equity, ΔCLV^bMExp = customer lifetime value before marketing expenditures, ΔLCA = lifetime acquisition expenditures, ΔLCR = lifetime retention expenditures, ΔTLCA = total lifetime acquisition expenditures, and ΔTLCR = total lifetime retention expenditures.

customer equity before marketing expenditures of existing customers ($45.01 million), lost customers (–$7.26 million), and new customers ($22.69 million).

Furthermore, Figure 6 indicates the changes in customer metrics, customer value metrics, and customer quantity metrics and thus summarizes what has happened during the period and the forward-looking effects of those changes (i.e., changes in customer equity). For example, investors might note that Netflix increased its existing customers (.49 million), lost more customers than in Q2 2006 (–.05 million), and gained more customers than it did in Q2 2006 (.24 million) and therefore increased the value of the whole customer base, primarily because its average customer cash flow ($.73) increased during that period.

In addition to decomposing changes in customer equity for several components, investors might want to know which metrics caused those changes, as we provide in Table 4, which includes the total effect (total change), value effects (changes due to changes in customer value metrics), quantity effects (changes due to the number of existing, lost, and new customers), and interaction effects (changes due to simultaneous changes in customer value and quantity metrics).

According to Table 4, the major sources of Netflix's increased customer equity in Q2–Q3 2006 ($48.10 million) are positive value ($22.35 million) and quantity ($23.22 million) effects. Furthermore, the change in customer cash flow increases customer equity by $21.98 million, supported by the increase in customer lifetime ($1.84 million) but partly compensated for by higher acquisition expenditures (–$1.47 million). In addition to this decomposition, management can explain why those changes occurred and what it plans to do in the future.

According to Netflix's financial statements, monthly revenues per subscriber declined because of the continued popularity of lower-cost subscription plans. However, the cost of subscription revenues (revenue-sharing expenses, amortization of the DVD library, and postage and packaging expenses) and fulfillment (expenses incurred in operating and staffing shipping and customer service centers, including receiving, inspecting, and warehousing the library and credit card fees) declined as a percentage of revenue.

TABLE 4
Netflix's Customer Flow Equity Statement (Q2–Q3 2006): Effects View

	Q1–Q2 2006	Q2–Q3 2006	Trend
Total Effect (In Thousands of Dollars)	63,543.91	48,098.98	▼
Value Effects	27,342.54	22,351.47	▼
Customer cash flow	28,515.11	21,977.91	▼
Customer lifetime	6,373.39	1,839.46	▼
Acquisition expenditures	–7,545.96	–1,465.90	▲
Quantity Effects	32,166.34	23,218.85	▼
Lost customers	–47,269.71	–56,503.34	▼
New customers	79,436.05	79,722.18	▲
Interaction Effects	4,035.03	2,528.67	▼
Lost customers	–5,664.43	–3,782.62	▲
Customer cash flow	–4,494.67	–3,473.78	▲
Customer lifetime	–1,169.75	–308.84	▲
New customers	8,949.88	6,179.70	▼
Customer cash flow	6,270.28	5,569.95	▼
Customer lifetime	2,679 60	609.75	▼
Lifetime			
Customer cash flow	1,753.39	234.67	▼
Other	–1,003.80	–103.08	▲

Netflix explains this decline with the lowered cost per paid shipment, which includes a drop in the percentage of DVDs mailed to subscribers subject to revenue-sharing agreements. The firm also notes the decline in overall usage (i.e., fewer monthly movie rentals per average subscriber) and increased operational efficiencies. These changes increase our calculated customer cash flow. Finally, the minor increase in retention rate appears to stem from the firm's price parity with Blockbuster, as well as service improvements (e.g., more titles, better recommendations).

Furthermore, the quantity effects in Q2–Q3 2006 are positive ($23.22 million), indicating that Netflix grows its customer base. The positive interaction effects ($2.53 million) also suggest that the cash flow and retention rate increases prompt positive customer cash flow and lifetime effects for existing and new customers but results in a negative effect for lost customers.

To enhance understandability, Table 4 also includes trends compared with the previous quarter. Thus, investors might observe changes in the customer equity trends over time. For example, a comparison of Q2–Q3 2006 with Q1–Q2 2006 (see Table 4) shows that Netflix's customer equity growth slows down. Although there are positive trends in acquisition expenditures and new customers, the trends of customer cash flows, customer lifetime, and lost customers are deteriorating. These trends allow investors to evaluate the firm's ability to solve previous periods' problems or its potential to outperform its previous growth in customer equity.

Discussion and Conclusion

We emphasize that the reporting of forward-looking customer metrics addresses the demand for additional information that facilitates investors' decisions. Therefore, we propose a means to report customer equity that matches financial reporting criteria and enables investors, creditors, and other consumers of financial reports to understand the firm's capability to generate shareholder value. In this sense, our research contributes to recent discussions about financial reporting and enables investors to monitor a firm's performance with respect to its primary assets. Moreover, it contributes to the discussion about marketing accountability and might support marketing's reentry into the boardroom because it aligns customer management with corporate goals and the investor's perspective.

For the specific purpose of reporting, we focus on the value of the current customer base and its changes over time. Customer-based firm valuation would be a natural extension, and the value of future customers must then be taken into account. A rather limited number of studies exist in this area. Kim, Mahajan, and Srivastava (1995) estimate the value per pop (i.e., the number of people living in a service area) using subscriber data in the cell phone industry. They find that their model is able to capture and predict this value well. Gupta, Lehmann, and Stuart (2004) use publicly available customer data from five firms to estimate customer equity. They find that their estimates of customer equity are reasonably close to the market value for three of the five firms (exceptions were Amazon.com and eBay). Libai, Muller, and Peres (2007) replicate their model using a Bass diffusion model with customer defection. Their results confirm Gupta, Lehmann, and Stuart's (2004) findings. Rust, Lemon, and Zeithaml (2004) use a survey of 100 airline customers to estimate CLV for American Airlines. Using this estimate and the total number of airline passengers, they estimate the overall customer value of American Airlines in 1999 as $7.3 billion. Because this estimate does not include international traffic and other nonflight sources of revenue, it is reasonably close to the $9.7 billion market value of American Airlines at that time. We believe that if more firms follow the example of companies such as Netflix and disclose better information about their customer base, models for calculating and decomposing the value of the customer base (i.e., customer equity) and customer-based firm valuation will gain more importance in practice.

The proposed technique for customer equity reporting permits a wide range of models for calculating CLV and customer equity. We introduce a parsimonious model that reflects our data availability for calculating and decomposing customer equity in application to data from Netflix. The provided calculations and decomposition of customer equity reveal a clear understanding of the health of Netflix's customer base to investors, especially because they also analyze the influence of changes in customer metrics on the value of the customer base. As such, our model is both diagnostic and forward looking.

Our research also contains limitations, which suggest opportunities for additional study. Although we develop our reporting technique for a broad range of firms, we study only a single firm in depth. Further research might analyze the value of the customer base of additional firms and industries to examine whether industry-specific patterns of changes emerge in terms of customer metrics. Such an analysis could also provide a benchmark for developments across not only time but also industries. In addition, our specific calculation of CLV entails just one possible method that requires only a limited amount of information, whereas the general framework also permits several other specifications. We explain why violations of the assumptions we need to make should lead only to minor bias in the results, but additional research could analyze our arguments in greater depth. Further research might use our ideas when other types of data are available. For example, although individual customer information is sensitive and firms might be reluctant to disclose them in financial reports, aggregated segment-level statistics (e.g., high, middle, and low CLVs or the variance of average CLVs and other customer metrics across segments) may lead to additional insights for investors. Thus, instead of using only customer averages—across all the firm's customers—when reporting CLV, acquisition or retention expenditures, and customer equity, firms might also report the variance of these averages (even on a segment level) to account for heterogeneity across customers. Moreover, because of the limited amount of data, our approach does not include competitive effects. Using a Markov switching-matrix approach to model CLV would enables us to design a model that contains both customer acquisition and retention in the context of brand switching (e.g., Rust, Lemon, and Zeithaml 2004). Further research might incorporate this, if the necessary data are available.

An extended approach could also incorporate varying future acquisition or retention expenditures as well as discount and retention rates because they might change (e.g., Gupta and Lehmann 2005). Likewise, our specification could be extended to capture additional aspects that could be incorporated as additional customer metrics, such as cross- and up-selling projections. Moreover, further research could examine how our technique should be adjusted for noncontractual relationships because we focus on firms with contractual relationships, which can easily determine the number of existing and lost customers at a particular point in time. Additional research might also examine implementation issues of our approach. For example, it would be worthwhile to know how to stimulate usage of such reporting techniques for a firm's investor-relations activities, which information they already compute and track over time, and which level of metrics they are willing to disclose. Thus, activity-based costing approaches might avoid possible endogeneity problems related to our way of computing acquisition or retention expenditures.

Future approaches could also incorporate additional customer performance metrics. For example, metrics of customer value creation and marketing efficiency may offer additional information on the firm's customer management activities over time. Examples are metrics that indicate the creation of value through new customer acquisition activities or demonstrate changes in the efficiency of investments in customer acquisition or retention activities over time. Further research might examine whether additional metrics provide better diagnostics of future firm performance than current metrics.

We propose a reporting technique that enables investors to analyze the situation of the firm better and to identify the metrics responsible for changes in customer equity. However, we do not intend to explain why certain metrics have changed; that job is up to management, as are explanations of future plans (IASB 2005). Our reporting technique visualizes the long-term effects of changes for external recipients of financial reports. Internally, firms should employ models that diagnose the reasons for particular changes in customer metrics in much more detail (e.g., Rust, Lemon, and Zeithaml 2004). Further research could explore the advantages of combining our reporting technique with such detailed internal models. A particular challenge might be to determine how these combinations best meet the criteria for financial reporting—namely, future orientation, decomposition, objectivity, comparability, simplicity, and cost effectiveness.

Summary

Recent discussions in accounting have demanded the need for additional information that facilitates investors' decision making and meets the objectives of financial reporting. Forward-looking customer metrics are necessary and useful as a managerial tool and thus should also be reported in financial statements to enable investors to understand clearly the firm's capability to generate shareholder value. An external reporting technique that addresses the consequences of a firm's customer management activities completes the concept of value-based customer management because it aligns customer management with corporate goals and the investor's perspective. We develop a reporting technique that provides a starting point for considering the value of the customer base in a firm's financial report, especially in the "MD&A" or "Management Commentary" sections. We emphasize that this reporting technique provides a supplement and complement to current information in financial statements, as demanded by the SEC and IASB. Because marketing literature includes in-depth discussions of the concepts of customer and brand equity, marketing academics should take a leading role in transferring that knowledge to other areas, such as accounting or finance.

REFERENCES

Ambler, Tim (2000), *Marketing and the Bottom Line*. London: Financial Times/Prentice Hall.

Amir, Eli and Baruch Lev (1996), "Value-Relevance of Nonfinancial Information: The Wireless Communication Industry," *Journal of Accounting and Economics*, 22 (1–3), 3–30.

Berger, Paul D. and Nada L. Nasr (1998), "Customer Lifetime Value: Marketing Models and Applications," *Journal of Interactive Marketing*, 12 (1), 18–30.

Blattberg, Robert C. and John Deighton (1996), "Managing Marketing by the Customer Equity Test," *Harvard Business Review*, 74 (July–August), 136–44.

Deloitte (2006), *Global Capital Markets and the Global Economy: A Vision from the CEOs of the International Audit Networks*. London: Deloitte.

Fader, Peter S. and Bruce G.S. Hardie (2006), "Customer-Base Valuation in a Contractual Setting: The Perils of Ignoring Heterogeneity," working paper, London Business School.

FASB (1978), *Statement of Financial Accounting Concepts No. 1: Objectives of Financial Reporting by Business Enterprises*. Stamford, CT: FASB.

——— (1980), *Statement of Financial Accounting Concepts No. 2: Qualitative Characteristics and Accounting Information*. Stamford, CT: FASB.

——— (2000), *Statement of Financial Accounting Concepts No. 7: Using Cash Flow Information and Present Value in Accounting Measurements*. Stamford, CT: FASB.

Fischer, Marc (2007), "Valuing Brand Assets: A Cost-Effective and Easy-to-Implement Measurement Approach," Marketing Science Institute Working Paper No. 07-107.

Fornell, Claes, Sunil Mithas, Forrest V. Morgeson, and M.S. Krishnan (2006), "Customer Satisfaction and Stock Prices: High Returns, Low Risk," *Journal of Marketing*, 70 (January), 3–14.

Gupta, Sunil and Donald R. Lehmann (2003), "Customers as Assets," *Journal of Interactive Marketing*, 17 (1), 9–24.

——— and ——— (2005), *Managing Customers as Assets*. Upper Saddle River, NJ: The Wharton School Publishing.

———, ———, and Jennifer A. Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (February), 7–18.

IASB (2004), *Framework for the Preparation and Presentation of Financial Statements*. London: IASB.

——— (2005), "Management Commentary." London: IASB.

Ittner, Christopher D. and David F. Larcker (1998), "Are Nonfinancial Measures Leading Indicators of Financial Performance? An Analysis of Customer Satisfaction," *Journal of Accounting Research*, 36 (3, Supplement), 1–35.

———, ———, and Taylor Randall (2003), "Performance Implications of Strategic Performance Measurement in Financial Service Firms," *Accounting, Organizations and Society*, 28 (7–8), 715–41.

Kim, Namwoon, Vijay Mahajan, and Rajendra K. Srivastava (1995), "Determining the Going Market Value of a Business in an Emerging Information Technology Industry: The Case of the Cellular Communications Industry," *Technological Forecasting and Social Change*, 49 (3), 257–79.

Kumar, V. and Morris George (2007), "Measuring and Maximizing Customer Equity: A Critical Analysis," *Journal of the Academy of Marketing Science*, 35 (2), 157–71.

———, Girish Ramani, and Timothy Bohling (2004), "Customer Lifetime Value Approaches and Best Practice Applications," *Journal of Interactive Marketing*, 18 (3), 60–72.

——— and Werner Reinartz (2005), *Customer Relationship Management: A Databased Approach*. New York: John Wiley & Sons.

Libai, Barak, Eitan Muller, and Renana Peres (2007), "The Influence of Attrition on the Growth and Equity of Competitive Services," Marketing Science Institute Working Paper No. 07-109.

Marketing Science Institute (1999), "Value of the Brand," workshop at Marketing Science Institute Conference on Marketing Metrics, Washington, DC (October 6–8).

McGovern, Gail J., David Court, John A. Quelch, and Blair Crawford (2004), "Bringing Customers into the Boardroom," *Harvard Business Review*, 82 (11), 1–10.

Nagar, Venky and Madhav V. Rajan (2005), "Measuring Customer Relationships: The Case of the Retail Banking Industry," *Management Science*, 51 (6), 904–919.

Netflix (2006), *SEC Filing 10-Q-Form*. Los Gatos, CA: Netflix.

Payne, Adrian and Pennie Frow (2005), "A Strategic Framework for Customer Relationship Management," *Journal of Marketing*, 69 (October), 167–76.

PricewaterhouseCoopers (2005), *Trends 2005: Good Practices in Corporate Reporting*. London: PricewaterhouseCoopers.

Reinartz, Werner J. and V. Kumar (2000), "On the Profitability of Long Lifetime Customers: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64 (October), 17–35.

——— and ——— (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (January), 77–99.

———, Jacquelyn S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," *Journal of Marketing*, 69 (January), 63–79.

Rust, Roland T., Katherine N. Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (January), 109–127.

Ryals, Lynette J. (2005), "Making Customer Relationship Management Work: The Measurement and Profitable Management of Customer Relationships," *Journal of Marketing*, 69 (October), 252–61.

Said, Amal A., Hassan R. HassabElnaby, and Benson Wier (2003), "An Empirical Investigation of the Performance Consequences of Nonfinancial Measures," *Journal of Management Accounting Research*, 15 (1), 193–223.

SEC (2003), "Commission Guidance Regarding Management's Discussion and Analysis of Financial Condition and Results of Operations," Release Nos. 33-8350; 33-48960, FR-72. New York: SEC.

Smith, Rodney E. and William F. Wright (2004), "Determinants of Customer Loyalty and Financial Performance," *Journal of Management Accounting Research*, 16 (1), 183–205.

Srinivasan, Shuba and Dominique M. Hanssens (2007), "Marketing and Firm Value: Metrics, Methods, Findings and Future Directions," working paper, Gary Anderson School of Management, University of California, Riverside.

Venkatesan, Rajkumar and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68 (October), 106–125.

Shuba Srinivasan, Koen Pauwels, & Vincent Nijs

Demand-Based Pricing Versus Past-Price Dependence: A Cost–Benefit Analysis

The authors develop a conceptual framework of the factors that motivate a retailer's decision to rely on demand conditions and past prices in setting current and future prices. Specifically, they examine the circumstances under which retailers choose demand-based pricing versus past-price dependence for different brands and categories. Given scarce resources and costs of price adjustments, demand-based pricing is more likely when the customer-driven and firm-driven costs of adjusting pricing patterns are low or when the benefits of such adjustments are high. First, the customer-driven benefits of demand-based pricing are expected to be greater in categories with higher penetration and for brands with higher market share and higher demand sensitivity to price. Second, the firm-driven benefits are greater for categories with higher private-label share. Finally, the customer-driven costs are greater for expensive categories, whereas the firm-driven costs are greater for categories with many stockkeeping units. The empirical findings support the conceptual framework, implying that customer-driven and firm-driven benefits are the main stimulants in the retailer's choice of demand-based pricing. In contrast, customer-driven and firm-driven costs significantly hinder retailer implementation of demand-based pricing. These insights enable retailers to identify problem areas and opportunities to improve the allocation of scarce pricing resources. The results also contribute to the ongoing debate in economics and marketing on the rationality of observed past-price dependence. Whereas previous research points to the negative impact on gross margins of this practice, the authors find that retailers weigh the costs and benefits of demand-based pricing rather than adhere to past-pricing patterns.

Keywords: demand-based pricing, past-price dependence, retail-price drivers, time-series models, generalized forecast error variance decomposition

Retailers face the complicated task of setting and changing prices for the many items they carry. A typical grocery store in the United States now carries more than 31,000 items in hundreds of product categories (Kahn and McAlister 1997). Apart from the sheer number of price change possibilities, the considerations that enter retailers' pricing decisions have become very complex. Sophisticated demand forecasts based on scanner data, the push toward category management, and marketing intelligence on competing retailers' prices are all important and have been incorporated in recent analytical research (e.g., Basuroy, Mantrala, and Walters 2001; Kim and Staelin 1999; Wedel, Zhang, and Feinberg 2004). However, empirical studies have found that retailers often choose not to adapt prices on the basis of demand conditions (Dutta, Bergen, and Levy 2002), leading to past-price dependence and lower category margins (Nijs, Srinivasan, and Pauwels 2007).

Although past-price dependence, or price rigidity, is underexplored in the marketing literature, it is a fundamental issue in pricing (Bergen et al. 2003).[1] Classical economic theory assumes that prices adjust flexibly in response to changes in demand and costs, and most research in marketing adopts this assumption, either directly or implicitly.[2] The alternative of complete price rigidity is at odds with the large variation in prices that, for example, Bils and Klenow (2004) and Gordon (1981) observe. Other schools of thought in economics, such as new Keynesian macroeco-

Shuba Srinivasan is Associate Professor of Marketing, A. Gary Anderson School of Management, University of California, Riverside (e-mail: shuba.srinivasan@ucr.edu). Koen Pauwels is Associate Professor of Business Administration, Tuck School of Business, Dartmouth College (e-mail: koen.h.pauwels@dartmouth.edu). Vincent Nijs is Assistant Professor of Marketing, Kellogg School of Management, Northwestern University (e-mail: v-nijs@kellogg.northwestern.edu). The authors are listed in reverse alphabetical order. They are grateful to the Dominick's project at the Graduate School of Business, University of Chicago, for making the data available. The article benefited from the comments of two anonymous *JM* reviewers and by seminar participants at the 2006 Marketing Science Conference at Pittsburgh, the Marketing Dynamics Conference at University of California, Los Angeles, University of Iowa, University of Groningen, Tilburg University, HEC Paris, University of California, Berkeley, and Washington University in St. Louis. The authors also thank Marnik Dekimpe and Gary Russell for their comments and suggestions.

---

[1] We thank an anonymous reviewer for constructive suggestions on the discussion of price rigidity.

[2] We view the term "pricing rigidity" as an indication that prices are not flexibly adjusted to changes in demand or costs. However, this does not imply that prices never change; rather, pricing patterns follow their usual shape, even in the presence of demand/cost changes. We expand on this definition in the "Methodology" section.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

nomic theory (e.g., Blinder 1991; Levy 2007), and work in industrial organization (e.g., Carlton 1986) entertain the possibility that prices are rigid to some degree. The debate of rigid versus flexible prices, which lies at the heart of the theories of firms, markets, industries, and economies (Golosov and Lucas 2007; Zbaracki et al. 2004), has been the subject of empirical and theoretical studies (for an extensive overview, see Wolman 2007). In marketing, however, only a handful of researchers have addressed the issue of price rigidity.

On the one hand, prices may exhibit rigidity because retailers suboptimally anchor their pricing decisions on the past (Krishna, Mela, and Urbany 2001) or simply lack detailed information about market demand (*BusinessWeek* 2000) and appropriate tools for making pricing decisions (AMR Research 2000). On the other hand, retailers may have good reasons to maintain consistent pricing patterns. These may include the high managerial and physical costs of considering and executing alternative pricing patterns (Levy et al. 1997; Slade 1998; Zbaracki et al. 2004), as well as legal, goodwill, and customer reference price issues linked to unexpected price fluctuations (Bergen et al. 2003).

Previous research has mainly focused on the general occurrence of past-price dependence, not on the circumstances that lead a given retailer to behave this way for some categories and brands but not for others, as Nijs, Srinivasan, and Pauwels (2007) report. Nijs, Srinivasan, and Pauwels quantify the relative importance of different drivers of retail prices in a large-scale empirical study. They show that retail prices are driven by, in order of importance, past prices, wholesale prices, brand demand, category management, and store traffic/interretailer price competition. They further demonstrate that the influence of these drivers on retailer pricing tactics varies greatly by category and brand and is linked to retailer category margins. Specifically, of the different drivers, demand-based pricing is most strongly associated with higher retailer margins. In contrast, the most influential price driver, past-price dependence, is linked to lower retailer margins.[3] To illustrate the importance of these effects, consider that the average retail margin in the Dominick's Finer Foods database is approximately $525 per category per store per week. From Nijs, Srinivasan, and Pauwels's (2007) analysis, we calculate that a 10% increase in the influence of past-price dependence on retail price setting will reduce weekly category margins by $23. However, this increased emphasis on past-price dependence also implies a reduction in the relative influence of another retail-price driver. If the increase in past-price dependence comes at the expense of demand-based pricing, there would be an additional negative margin impact of $177. The net impact of these two forces would result in a drop in margins of $200 (38%).

Given the theoretical and monetary importance of this phenomenon, "it is unfortunate that so little attention has been given to characterizing the circumstances that give rise to high versus low nominal levels of price inertia" (Andrew Caplin, qtd. in Levy et al. 1998, p. 81). We contribute to this field of inquiry by integrating relevant theories and empirically testing some of their implications. Our key research question is, Under which conditions do retailers rely more heavily on demand-based pricing than past-price dependence in setting prices? Thus, we aim to increase the understanding of demand-based pricing and past-price dependence with an empirical investigation into the variation in both practices across brand and categories.

To this end, we develop a conceptual framework of the cost versus benefit trade-offs between demand-based pricing and past-price dependence in the next section. We then introduce the methodology and report the results of our analysis. We conclude with managerial implications, contributions, and areas for further research.

Conceptual Framework

Previous marketing research has examined sources of price variation from both the manufacturer and the retailer perspective. For manufacturers, Raju, Srinivasan, and Lal (1990) show that brands with lower loyalty have more to gain from promotions, and Kinberg, Rao, and Shakun (1974), Lal (1990), and Rao (1991) argue that promotions by premium brands can keep an intruder, such as a private label, from encroaching on their customers. Retailers may vary prices because of decreasing unit variable costs (Blattberg and Neslin 1990) or a desire to transfer holding costs to consumers (Blattberg, Eppen, and Lieberman 1981). In addition, Varian (1980) shows that retailers may implement sales to price discriminate between informed and uninformed consumers, and Kopalle, Rao, and Assunção (1996) generalize Greenleaf's (1995) result that reference price formation may induce the retailer to vary prices over time. Indeed, when enough consumers weigh price gains more than price losses, the optimal pricing policy is high–low (Kopalle, Rao, and Assunção 1996). Implementing such a pricing policy would require retailers to develop and maintain a thorough understanding of consumers across a multitude of categories. Fader and Lodish (1990, p. 55) argue that retailers are unlikely to go to such lengths, attributing a lack of promotional activity to the observation that many categories "are 'unglamorous' and thus receive no special attention from retailers." They identify determinants of category promotional activity using data from Information Resources Inc.'s (1997) *Marketing Factbook*, but they do not address the extent to which retail pricing in a category is driven by demand-based pricing versus past-price dependence.[4]

---

[3] We define "past-price dependence" as the extent to which variability in retail prices over time can be attributed to past fluctuations in a brand's own price. Past-price dependence can result in everyday same prices or high–low pricing, but without regard for the state of demand or costs. We define the extent of "demand-based pricing" as the variability in retail prices that is attributable to fluctuations in a brand's past sales volume. We provide details in the "Methodology" section and in Technical Appendix A.

[4] Pricing for a category with high promotional activity can be driven by either demand-based pricing or past-price dependence. If the occurrence of promotions follows from changes in consumer demand, demand-based pricing will be more prominent. Conversely, if the pattern of promotions can be predicted from past pricing patterns rather than changes in demand, high past-price dependence will be observed.

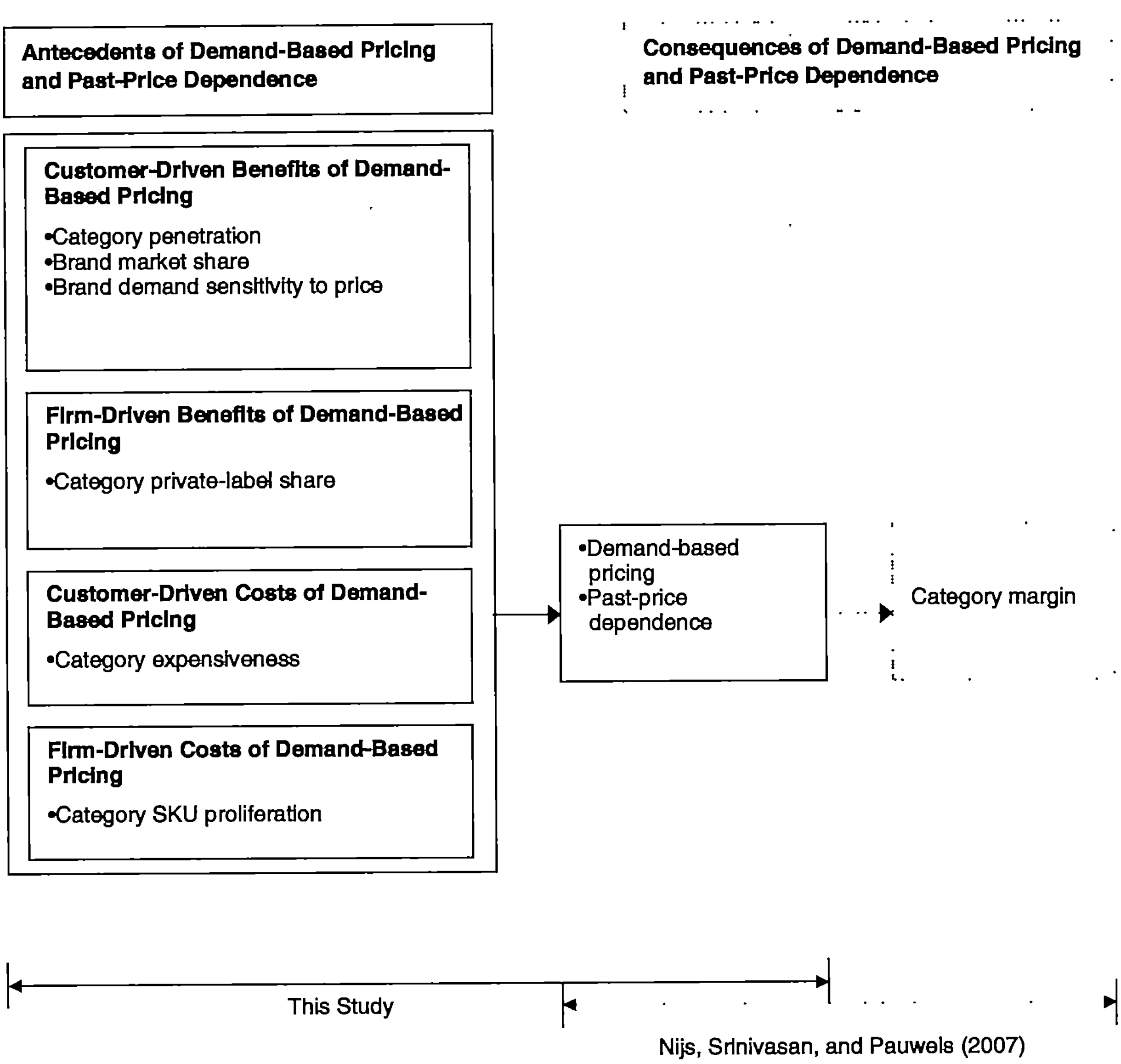
Our framework goes beyond the arguments, theories, and variables used in previous research and attempts to explain when and why retailers choose to engage in demand-based pricing versus past-price dependence. Our basic premise is that retailers make a cost–benefit trade-off when deciding whether to rely on demand-based pricing. Although they may not all be directly observable to the researcher, we argue that inferences can be drawn about the nature of these costs and benefits from observable variables.

We distinguish two types of costs and benefits of demand-based pricing: firm driven and customer driven. We consider the material, managerial, and labor costs as well as the margin benefits of pushing the retailer's private label to be firm driven. Customer-driven costs and benefits, which refer to customers' reactions to changes in pricing patterns, can include purchase behavior, reference price formation, and the customer's perception of the retailer. Our conceptual framework integrates each of these forces to provide a consistent description of how retailers trade off the costs and benefits of demand-based pricing (see Figure 1).[5]

Customer-Driven Benefits of Demand-Based Pricing

Effective pricing requires a retailer to allocate scarce pricing resources for the largest returns. We expect the retailer to do so in accordance with the perceived importance of the category and brand to the retailer's performance objectives. In particular, we propose that the benefits of demand-based price adjustments are larger in categories with higher penetration and for brands with high market share and demand

FIGURE 1
Conceptual Framework


Antecedents of Demand-Based Pricing and Past-Price Dependence
Consequences of Demand-Based Pricing and Past-Price Dependence
Customer-Driven Benefits of Demand-Based Pricing
•Category penetration
•Brand market share
•Brand demand sensitivity to price
Firm-Driven Benefits of Demand-Based Pricing
•Category private-label share
Customer-Driven Costs of Demand-Based Pricing
•Category expensiveness
Firm-Driven Costs of Demand-Based Pricing
•Category SKU proliferation
•Demand-based pricing
•Past-price dependence
Category margin
This Study
Nijs, Srinivasan, and Pauwels (2007)

---

[5] Figure 1 also highlights the contribution of the current study over that of Nijs, Srinivasan, and Pauwels (2007). Whereas Nijs, Srinivasan, and Pauwels focus on the consequences of demand-based pricing and past-price dependence, we focus on their antecedents.

sensitivity to price. The more pronounced the benefits of demand-based pricing, the more this price driver should override the retailer's use of past-price dependence.

Category penetration. A key issue in pricing is that a retailer must decide on the role of each category in the overall store portfolio (Dhar, Hoch, and Kumar 2001). As McAlister (2008) argues, retailers have scarce resources to engage in demand-based pricing across hundreds of categories and thousands of products. The larger the proportion of households that purchase in the category (i.e., category penetration), the larger is the expected customer purchase reaction, and thus the larger are the revenue benefits of demand-based pricing (Fader and Lodish 1990). Indeed, prior research has suggested that category penetration is the most informative measure of category promotional activity (ibid). Thus, on the basis of the customer benefits of demand-based price adjustments, we expect the following:

> H_1: In setting prices of brands in categories with higher penetration rates, retailers place (a) greater emphasis on demand-based price adjustments and (b) less emphasis on past-price dependence.

Brand market share. Both analytical models (e.g., Lal, Little, and Villas-Boas 1996) and empirical evidence (Chevalier and Curhan 1976; Pauwels 2007; Walters 1989) suggest that retailers are more willing to deviate from established pricing patterns for high-share brands than for smaller brands. Leading brands enjoy greater consumer awareness and familiarity (Keller 1993), thus creating a larger customer base that may be affected by the retailer's changes to pricing patterns. Indeed, promotions on leading brands have the power to expand the category (Bronnenberg and Mahajan 2001) and even increase store traffic (Moorthy 2005).

> H_2: In setting prices of brands in categories with higher brand market share, retailers place (a) greater emphasis on demand-based pricing and (b) less emphasis on past prices.

Brand demand sensitivity to price. The benefits to demand-based pricing should be higher in categories in which demand is sensitive to changes in price. The retailer is then likely to alter pricing patterns according to perceived differences in consumer preferences and willingness to pay across brands and categories (Levy et al. 1998) and to change prices following shocks to demand.

> H_3: In setting retail prices of brands with higher demand sensitivity, retailers place (a) greater emphasis on demand-based pricing and (b) less emphasis on past prices.

Firm-Driven Benefits of Demand-Based Pricing

Category private-label share. An important part of a retailer's business is the private-label program. Some retailers (e.g., Wegman's) successfully use store brands as a source of differentiation and a revenue driver (Dhar, Hoch, and Kumar 2001). For others, the private label offers increased category profits as a result of higher percentage margins and increased bargaining power compared with national brand manufacturers (e.g., Pauwels and Srinivasan 2004). Because the retailer reaps the full rewards from private-label performance, categories with a higher retailer private-label share tend to get more pricing attention. Therefore, we expect more extensive use of demand-based pricing when firm-driven benefits are higher and less dependent on past prices.

> H_4: In setting prices of brands in categories with higher private-label share, retailers place (a) greater emphasis on demand-based pricing and (b) less emphasis on past prices.

Customer-Driven Costs of Demand-Based Pricing

Previous research has argued that the customer-driven costs of pricing adjustments are larger than the firm-driven costs and, thus, more important (Zbaracki et al. 2004). In Rotemberg's (2002) model, a threat of consumers' angry reactions over unfair price increases can lead to price rigidity. In line with this finding, Blinder and colleagues (1998) conclude that firms are less willing to engage in unanticipated changes to price because doing so would antagonize their customers. The more pronounced the customer costs of demand-based pricing, the less we expect retailers to opt for this price driver.

Category expensiveness. Customers are more likely to care about pricing adjustments in expensive categories because such categories consume a larger part of their budgets (Mazumdar and Papatla 2000). Even when justified on the basis of overall demand considerations, some customers may consider these pricing adjustments "unfair," either because they will need to pay more or because they recently paid more for a product that is now cheaper. Retailers may also hesitate to cut prices because "customers will misinterpret the price cuts as reductions in quality" (Blinder et al. 1998, p. 173). Therefore, retailers may adhere to well-established pricing patterns with a consistent pattern of discounts for expensive categories (Fader and Lodish 1990). Thus, the observed past-price dependence may be the result of careful consideration of the long-term pros and cons of consistency in pricing patterns.

> H_5: In setting prices of brands in expensive categories, retailers place (a) less emphasis on demand-based pricing and (b) greater emphasis on past prices.

Firm-Driven Costs of Demand-Based Pricing

The firm-driven costs of demand-based pricing have only recently been the subject of academic inquiry. Most important are the management time and attention required to gather the relevant information and to make and implement pricing decisions (Zbaracki et al. 2004). The more pronounced the firm-driven costs of demand-based pricing, the less we expect retailers to rely on this price driver.

Category stockkeeping unit proliferation. A large number of stockkeeping units (SKUs) in a category makes it costly for retailers to evaluate alternative pricing schemes at a weekly level (McAlister 2008). For example, in a Dominick's store in the Chicago area, the oatmeal category has 96 SKUs, whereas shampoo has more than 2500 SKUs. Given the required effort and the pricing complexity in the latter category, we expect the retailer to engage less in demand-based pricing.

H_6: In setting prices of brands in categories with a high degree of SKU proliferation, retailers place (a) less emphasis on demand-based pricing and (b) greater emphasis on past prices.

Table 1 presents an overview of our hypotheses.

Methodology

We obtained estimates of demand-based pricing and past-price dependence, based on generalized forecast error variance decomposition (GFEVD), from Nijs, Srinivasan, and Pauwels (2007). These authors use weekly store-level scanner data from the Dominick's retail chain for 24 product categories in 85 stores. The results are available for the top three brands in each category in each store. In essence, GFEVD quantifies the relative influence on a brand's retail price variation over time of shocks that can be attributed to contemporaneous and past changes in each of the endogenous variables in the vector autoregressive model with exogenous variables (VARX), including brand demand, brand price, brand cost (wholesale price), demand and wholesale prices for competing brands (category management), and store traffic. Although store-level retail scanner data have been used in marketing to study pricing, they have been "rarely used by economists" (Levy, Dutta, and Begen 2002, p. 202) to study price rigidity (cf. Müller et al. 2006). For estimation details, see Technical Appendix A.

As we described previously, our focus is on two output metrics generated by GFEVD: DBP_{ijk} measures the extent to which current and past changes in demand for brand i in category k in store j drive prices for that same brand (i.e., demand-based pricing), and PPD_{ijk} measures the extent to which past prices for brand i in category k in store j drive prices for that same brand (i.e., past-price dependence). Demand-based pricing and past-price dependence account for 11.4% and 49.6% of the dynamic variation in retail prices, respectively. We report summary statistics per category in Table 2 (for an extensive discussion of other price drivers, see Nijs, Srinivasan, and Pauwels 2007).

We test our hypotheses by linking the estimates of demand-based pricing and past-price dependence to the customer- and firm-driven costs and benefits identified in our conceptual framework. Measurement details on these variables appear in Table 3.

We estimate the following equations:

$$(1)\quad DBP_{ijk} = \gamma_0 + \gamma_1 CPEN_k + \gamma_2 BMS_{ijk} + \gamma_3 BDSP_{ijk} + \gamma_4 CPLS_{jk} + \gamma_5 CE_{jk} + \gamma_6 CSKUP_{jk} + \mu_{ijk}, \text{ and}$$

$$(2)\quad PPD_{ijk} = \beta_0 + \beta_1 CPEN_k + \beta_2 BMS_{ijk} + \beta_3 BDSP_{ijk} + \beta_4 CPLS_{jk} + \beta_5 CE_{jk} + \beta_6 CSKUP_{jk} + \varepsilon_{ijk},$$

where DBP_{ijk} and PPD_{ijk} are as defined previously, μ_{ijk} is the error term for Equation 1, and ε_{ijk} is the error term for Equation 2. The covariates are CPEN (category penetration), BMS (brand market share), BDSP (brand demand sensitivity to price), CPLS (category private-label share), CE (category expensiveness), and CSKUP (category SKU proliferation).[6] We allow for store fixed effects in both equations.

Estimation of Equations 1 and 2 by ordinary least squares (OLS) provides consistent parameter estimates (see Murphy and Topel 1985). However, parameter standard errors may be biased because demand-based pricing and past-price dependence are estimated quantities. We use a

TABLE 1
Overview of Hypotheses

Benefit and Cost Factors	Hypotheses	Category and Brand Characteristics	Effect on Demand-Based Pricing	Effect on Past-Price Dependence
Customer-driven benefits of demand-based pricing	H_1	Category penetration	+	–
	H_2	Brand market share	+	–
	H_3	Brand demand sensitivity to price	+	–
Firm-driven benefits of demand-based pricing	H_4	Category private-label share	+	–
Customer-driven costs of demand-based pricing	H_5	Category expensiveness	–	+
Firm-driven costs of demand-based pricing	H_6	Category SKU proliferation	–	+

---

[6] We acknowledge that several other category and brand characteristics have been examined as moderators of retail price effects on performance (e.g., Srinivasan et al. 2004). These include product storability, category purchase frequency, category growth rate, and manufacturer promotional depth and frequency. Although we do not have a formal expectation of the direction of their effects, we can include them (separately or all together) in Equations 1 and 2 to control for their potential influence. Doing so does not affect our substantive findings, so we focus on the stated equations for clarity and conciseness.

TABLE 2
Extent of Demand-Based Pricing and Past-Price Dependence Across Categories

	Demand-Based Pricing (%)			Past-Price Dependence (%)		
Categories	25th	Mdn	75th	25th	Mdn	75th
Analgesics	7.6	10.7	12.8	47.1	48.7	50.7
Beer	11.3	13.8	17.4	48.2	52.2	57.9
Bottled juice	7.7	16.2	19.9	45.7	51.4	57.5
Canned soup	10.0	16.9	20.8	43.3	49.5	53.2
Cereal	16.6	19.9	22.5	44.3	48.0	53.3
Cheese	11.0	17.0	20.3	40.6	45.2	49.9
Cookies	7.2	21.5	24.8	40.8	42.7	47.4
Crackers	8.1	13.5	16.9	40.3	45.8	55.0
Dish detergent	2.2	5.6	10.4	54.7	59.7	64.3
Fabric softeners	5.2	7.7	11.1	51.7	54.5	57.6
Front-end candles	2.4	5.6	9.7	43.1	51.7	56.9
Frozen juice	10.4	13.9	17.9	50.6	55.2	58.8
Laundry detergent	6.6	8.8	13.1	44.9	48.5	51.4
Oatmeal	6.5	9.6	18.7	40.9	45.8	48.8
Paper towels	11.4	15.7	19.6	38.4	41.6	52.3
Refrigerated juice	15.2	18.5	21.0	43.9	49.2	54.5
Shampoos	1.7	3.7	5.8	46.1	52.5	60.1
Snack crackers	9.2	11.8	14.7	46.5	51.5	59.3
Soap	2.3	4.2	9.8	44.2	48.9	52.0
Soft drinks	2.7	5.4	24.3	45.6	47.2	49.0
Toilet tissue	12.3	19.0	22.2	39.4	43.8	53.1
Toothbrush	2.7	4.3	6.7	54.2	56.7	59.7
Toothpaste	2.5	9.1	13.6	49.9	52.7	59.0
Tuna	9.4	12.4	15.5	40.7	43.7	46.4
Total	6.1	11.4	17.3	44.4	49.6	55.7

Notes: The remainder of retail price variation is accounted for by store traffic, wholesale price, and category management considerations, as detailed by Nijs, Srinivasan, and Pauwels (2007).

bootstrap algorithm to obtain unbiased standard errors (for details, see Technical Appendix B).

Results[7]

Customer-Driven Benefits of Demand-Based Pricing

Category penetration. As the standardized coefficients in Table 4 show, in categories with a higher rate of penetration, the extent of demand-based pricing is more pronounced, whereas the extent of past-price dependence is lower, consistent with H_{1a} ($\gamma_1 = .377, p < .01$) and H_{1b} ($\beta_1 = -.241, p < .01$). Among all considered variables, category penetration is the strongest detractor of past-price dependence and the strongest enhancer of demand-based pricing.

Brand market share. In setting retail prices of high-share brands, retailers place greater emphasis on demand-based pricing and less emphasis on past prices, consistent with H_{2a} ($\gamma_2 = .241, p < .01$) and H_{2b} ($\beta_2 = -.181, p < .01$).

Brand demand sensitivity to price. Brands with high demand sensitivity elicit more demand-based pricing and less past-price dependence, consistent with H_{3a} ($\gamma_3 = .181, p < .01$) and H_{3b} ($\beta_3 = -.125, p < .01$).

Firm-Driven Benefits of Demand-Based Pricing

Category private-label share. In support of H_{4a} ($\gamma_4 = .136, p < .01$) and H_{4b} ($\beta_4 = -.135, p < .01$), retailers place greater emphasis on demand-based pricing and less emphasis on past prices in categories in which the retailer's private label commands a larger share.

Customer-Driven Costs of Demand-Based Pricing

Category expensiveness. In setting the prices of brands in expensive categories, retailers place less emphasis on demand-based pricing and greater emphasis on past prices, in support of H_{5a} ($\gamma_5 = -.050, p < .05$). Although the effect on past-price dependence is also in the expected direction, it does not reach traditional significance levels ($\beta_5 = .044, p = .11$).

Firm-Driven Costs of Demand-Based Pricing

Category SKU proliferation. In categories with a large number of SKUs, retailers are less likely to use demand-based pricing, consistent with H_{6a} ($\gamma_6 = -.060, p < .01$). Of the variables considered, category SKU proliferation most strongly hinders demand-based pricing. In contrast to H_{6b}, we find that retailers also rely less on past prices in categories with many SKUs. This might suggest that they use other tactics to simplify pricing in these categories, such as cost-plus and/or category management–based pricing considerations. Overall, our empirical findings support the hypotheses (see Table 5).

---

[7] We generate the results using Ox Version 4.00 (see Doornik 2002).

TABLE 3
Variable Operationalization

Benefit and Cost Factors	Operationalization
Customer-Driven Benefits of Demand-Based Pricing	
Category penetration	Category penetration is operationalized as the percentage of households that buy this category, obtained from the *Marketing Fact Book* (Information Resources Inc. 1997)
Brand market share	The brand's market share is operationalized as the average volume-based share of the brand as in Srinivasan and colleagues (2004).
Brand demand sensitivity to price	A brand's demand sensitivity to price is defined as the percentage change in brand sales as a result of a price promotion of 1%.
Firm-Driven Benefits of Demand-Based Pricing	
Category private-label share	This variable is operationalized as the percentage share of the retailer private-label brands in the category.
Customer-Driven Costs of Demand-Based Pricing	
Category expensiveness	Following Raju (1992), we first compute the regular price (highest price over the data period) of each brand. The category-level expensiveness is calculated by the market share weighted average of the regular prices of the brands in the category.
Firm-Driven Costs of Demand-Based Pricing	
Category SKU proliferation	The number of SKUs in the category (Narasimhan, Neslin, and Sen 1996) captures the extent of SKU proliferation.

Managerial Implications and Conclusions

Because considering and executing demand-based price changes are costly, the benefits of doing so should outweigh the costs. In this article, we developed and tested a conceptual framework that outlines the cost–benefit trade-off motivating retailers' choice of demand-based pricing versus past-price dependence.

First, the customer-driven benefits of demand-based pricing are higher for categories with higher penetration, as well as for brands with high market share and high demand sensitivity to price. Because categories with high penetration (e.g., cereals, soft drinks) offer more customer-driven benefits than those with low penetration (e.g., fabric softeners, toothbrushes), they enjoy greater retailer focus on demand-based pricing and less emphasis on past prices. Second, we find that the firm-driven benefits of demand-based pricing are more pronounced for categories with higher private-label share, leading to less emphasis on past-price dependence. Finally, the customer-driven costs of demand-based pricing are higher for expensive categories, whereas the firm-driven costs are higher for categories with many SKUs. The relative importance of these effects is as follows:

- Category penetration is the strongest positive determinant of demand-based pricing, followed by brand market share. Conversely, category SKU proliferation most severely complicates the implementation of demand-based pricing.
- Category penetration, private-label share, and SKU proliferation all substantially lower the prominence of past-price dependence. The same applies for the brand-level measures: market share and demand sensitivity to price.

Implications for Manufacturers

Our findings may help manufacturers develop a deeper appreciation of retailers' benefits versus costs of demand-based pricing and past-price dependence. Specifically, the estimates suggest scenarios in which pass-through of trade deals is difficult to achieve because of the retailer's reliance on past prices. For example, smaller brands face such pass-through jeopardy because past-price dependence is much stronger for these brands, consistent with Pauwels (2007). Moreover, retailers are less likely to apply demand-based pricing in expensive categories, which represents a hurdle for manufacturers trying to convince retailers to change their established pricing patterns.

Our findings also identify areas in which manufacturers can provide pricing support to the retailer. For categories and brands in which the cost–benefit trade-off does not favor demand-based pricing, manufacturers can work to limit the costs or enhance the benefits. For example, category captains could support retailer pricing policies for categories with many SKUs in a way that leads to a win–win situation for both parties.

Implications for Retailers

Recent research has shown that demand-based pricing is associated with higher retailer gross margins, whereas past-price dependence is associated with lower retailer gross margins (Nijs, Srinivasan, and Pauwels 2007). In the current article, we investigate the conditions under which retailers choose to rely more on the heavy machinery of demand-based pricing than simply sticking with past-pricing patterns by developing and testing a conceptual framework of the customer- and firm-driven benefits and costs of demand-based pricing.

Our framework and findings offer retailers a way to review the allocation of pricing resources systematically across categories according to expected costs and benefits. For example, we show that in categories with high penetration, there are significant benefits to demand-based pricing. A retailer could evaluate whether categories with higher penetration in its stores are getting more attention in terms of demand-based pricing. If some are not, the retailer might want to evaluate whether there are specific costs to demand-based pricing that apply to that category, such as those linked to high SKU proliferation. Moreover, the retailer could choose to build pricing capabilities and/or lower the firm-driven costs of demand-based pricing (e.g., use a better decision support system to support pricing in categories

TABLE 4
Customer- and Firm-Driven Benefits and Costs of Demand-Based Pricing

	Demand-Based Pricing ($R^2 = .30$)		Past-Price Dependence ($R^2 = .16$)	
Benefit and Cost Factors	**Standardized Coefficients**	**SE**	**Standardized Coefficients**	**SE**
Customer-Driven Benefits of Demand-Based Pricing				
Category penetration	.377**	.019	–.241**	.024
Brand market share	.241**	.025	–.181**	.027
Brand demand sensitivity to price	.181**	.024	–.125**	.025
Firm-Driven Benefits of Demand-Based Pricing				
Category private-label share	.136**	.017	–.135**	.022
Customer-Driven Costs of Demand-Based Pricing				
Category expensiveness	–.050*	.021	.044	.027
Firm-Driven Costs of Demand-Based Pricing				
Category SKU proliferation	–.060**	.023	–.105**	.024

*Significant at the 5% level.
**Significant at the 1% level.
Notes: n = 5190. Parameters are standardized Store-specific intercepts are not shown for space considerations.

TABLE 5
Overview of Empirical Support for Conceptual Framework

Hypotheses	Demand-Based Pricing	Past-Price Dependence
Customer-Driven Benefits of Demand-Based Pricing		
H_1: Category penetration	√	√
H_2: Brand market share	√	√
H_3: Brand demand sensitivity to price	√	√
Firm-Driven Benefits of Demand-Based Pricing		
H_4: Category private-label share	√	√
Customer-Driven Costs of Demand-Based Pricing		
H_5: Category expensiveness	√	
Firm-Driven Costs of Demand-Based Pricing		
H_6: Category SKU proliferation	√	

with many SKUs). The retailer could also attempt to control and/or limit costs of demand-based pricing from the customer's perspective. For example, price guarantees could reduce consumer irritation due to changes in established pricing patterns for expensive items.

In Table 6, using estimates from Nijs, Srinivasan, and Pauwels (2007), we show the gross margin impact of changing the level of demand-based pricing (past-price dependence) to the highest (lowest) average level observed in our data. For example, for the analgesics category, the level of demand-based pricing is 10.70%, whereas the score for past-price dependence is 48.70%. If the retailer were to boost demand-based pricing in this category to 21.50%, the average margin for the category would go up by $191.01. If the level of past-price dependence could be reduced to 41.60%, margins would increase by $16.14. The combined gross margin impact of both changes comes to $207.14. Using in-house cost information, retailers can reevaluate the cost effectiveness of their pricing approaches for different categories and brands. As the example for the analgesics category shows, the retailer has the potential to increase gross category margins by up to 67.33% if the firm and/or customer cost of the altered pricing tactics can be kept low.

If we assume that the mix of price drivers underlying the retailer's pricing patterns is already optimal given its current pricing tools and management practices, the numbers in Table 6 can also be interpreted as cost measures of demand-based pricing (e.g., Slade 1998). Quantifying how substantial these costs are and how they vary across categories informs retailers about consumer and organizational issues that are central to their ability to evaluate the costs and benefits of demand-based pricing versus past-price dependence. As such, our findings give retailers the ability to deal better with these costs and their implications for a variety of decisions, ranging from setting and adjusting prices to interpreting competitive pricing actions to developing improved pricing processes and capabilities.

Contributions to Theory and Empirics

We believe that a central contribution of this article is to the literature on price rigidity and flexibility (Zbaracki et al. 2004). Nijs, Srinivasan, and Pauwels (2007) provide empirical evidence that, at least in retailing, prices are not perfectly flexible. The current research expands on this insight by enhancing the understanding of variation in the patterns of price adjustment. We identify a set of measurable conditions that can enhance or diminish the flexibility of prices to demand shocks. We believe that these insights can be of value in academic research on pricing in both the marketing and the economics literature, offering promising opportuni-

TABLE 6
Margin Impact of Changing the Level of Demand-Based Pricing (Past-Price Dependence) to the Highest (Lowest) Average Level Observed in the Data

Category	DBP (%)	PPD (%)	Impact DBP ($)	Impact PPD ($)	Total Impact ($)	Average Margin ($)	% Margin Increase
Analgesics	10.70	48.70	191.01	16.14	207.14	307.67	67.33
Beer	13.80	52.20	136.18	24.09	160.27	329.04	48.71
Bottled juice	16.20	51.40	93.74	22.27	116.01	764.70	15.17
Canned soup	16.90	49.50	81.36	17.96	99.31	562.73	17.65
Cereal	19.90	48.00	28.30	14.55	42.84	1,185.26	3.61
Cheese	17.00	45.20	79.59	8.18	87.77	2,092.68	4.19
Cookies	21.50	42.70	.00	2.50	2.50	902.87	.28
Crackers	13.50	45.80	141.49	9.55	151.03	275.30	54.86
Dish detergent	5.60	59.70	281.21	41.14	322.34	254.77	126.53
Fabric softeners	7.70	54.50	244.07	29.32	273.38	89.98	303.83
Front-end candies	5.60	51.70	281.21	22.96	304.16	158.74	191.61
Frozen juice	13.90	55.20	134.41	30.91	165.32	620.71	26.63
Laundry detergent	8.80	48.50	224.61	15.68	240.29	567.13	42.37
Oatmeal	9.60	45.80	210.46	9.55	220.01	193.55	113.67
Paper towels	15.70	41.60	102.58	.00	102.58	234.12	43.81
Refrigerated juice	18.50	49.20	53.06	17.27	70.33	1,059.77	6.64
Shampoos	3.70	52.50	314.81	24.77	339.58	324.52	104.64
Snack crackers	11.80	51.50	171.55	22.50	194.05	490.16	39.59
Soap	4.20	48.90	305.97	16.59	322.56	238.73	135.11
Soft drinks	5.40	47.20	284.74	12.73	297.47	961.45	30.94
Toilet tissue	19.00	43.80	44.21	5.00	49.21	334.76	14.70
Toothbrush	4.30	56.70	304.20	34.32	338.52	92.72	365.11
Toothpaste	9.10	52.70	219.31	25.23	244.53	176.93	138.21
Tuna	12.40	43.70	160.94	4.77	165.71	405.03	40.91

Notes. DBP = demand-based pricing. PPD = past-price dependence. The "Impact DBP" and "Impact PPD" columns show the change in "Average Margin" as a result of improving the DBP and PPD levels in a category to the highest average DBP and lowest average PPD levels reported by Nijs, Srinivasan and Pauwels (2007)—21.50% and 42.70%, respectively. "Average Margin" is the weekly gross margin for a category averaged across stores in the Dominick's database. Caveats: The "Total Impact" and "% Margin Increase" measures should be interpreted with caution The underlying assumption is that all model parameters used to generate the reported GFEVD estimates remain constant. The margin impact values can be used as approximate benchmarks of the costs of adjusting pricing to a higher DBP level and/or as the basis for a ranking of the relative potential for profit enhancement across categories.

ties for fundamental contributions to price theory and interdisciplinary research.

The results can be linked to several theories of price rigidity that Blinder and colleagues (1998) summarize and test. Two of these theories (cost-based pricing and constant marginal cost) suggest that rigidity in prices can be caused by cost processes. We can rule out these cost-based explanations as causes of past-price dependence as measured in this study. By including a measure of cost in the VARX model described in Technical Appendix A, we ensure that the stickiness observed in retail pricing patterns is not due to wholesale prices.

Our conceptual model can be linked to the theory of judging quality by price. Price stickiness can result from retailer hesitation to cut prices because "customers will misinterpret the price cuts as reductions in quality" (Blinder et al. 1998, p. 173). We argue that for expensive categories, retailers are likely to adhere to well-established pricing patterns with a consistent pattern of discounts. This theory provides a partial explanation for the strong past-price dependence we observe in these categories.

The strongest link exists with theories of costly price adjustment (Mankiw 1985). Zbaracki and colleagues (2004) enrich this theory by quantifying menu, managerial, and customer costs of price changes in an industrial market. Although we study stickiness in pricing patterns rather than stickiness in prices, our study provides strong support for the theory that prices often do not adjust flexibly or completely to cost, competitive, or demand shocks. In contrast, our study offers little insight into several other theories of price rigidity (e.g., coordination failure, hierarchical delays, implicit and nominal contracts). Empirical work that studies these theories with appropriate data is an important area for further research.

We also provide insights into the ongoing debate in economics and marketing on the rationality of observed past-price dependence. Whereas Krishna, Mela, and Urbany (2001) and Nijs, Srinivasan, and Pauwels (2007) point to the negative profit impact of this practice, our findings indicate that retailers may well be rationally weighing the costs and benefits of demand-based pricing versus past-price dependence (Carlton 1986). Under the assumption that retailers are able to select the mix of price drivers optimally, we quantify the lower bounds of the costs of demand-based pricing in Table 6.

Our empirical findings are consistent with those of Bils and Klenow (2004), who find that prices vary more for products with more elastic demand and in less expensive categories. The latter result is an unexplained surprise, but it is explained in our cost–benefit framework.

A growing body of research has focused on explaining cross-category differences in price elasticities (e.g., Bell, Chiang, and Padmanabhan 1999; Bolton 1989; Narasimhan, Neslin, and Sen 1996) and promotional activity (Fader and Lodish 1990). Our study complements this line of empirical research by investigating the firm- and customer-driven costs and benefits that determine the retailer's relative use of demand-based pricing versus past-price dependence.

Areas for Further Research

Our research suggests several valuable directions for both theoretical and empirical research. First, given that extant theory focuses on the analysis of single-product settings with few exceptions (e.g., Lach and Tsiddon 1996), it might be fruitful to expand these theories to incorporate benefit and costs dimensions of price adjustments in a multiproduct setting (e.g., Midrigan 2006). Second, research using data from diverse organizational/retailer contexts could also test for the hierarchy theory that organizational inertia is a driver of price rigidity (Blinder et al. 1998). Third, such data sets on heterogeneous firms could also be helpful in testing for the price coordination theory that price leaders (versus price followers) drive patterns of price adjustments (ibid). Fourth, data on promotional calendars set at the start of the year and any adjustments made thereafter might provide insights into the importance of nominal contracts in past-price dependence (ibid). Fifth, whereas we focus on the firm-driven benefits and costs of demand-based pricing from the retailer perspective, it would be desirable to expand this theory to the supply side (i.e., wholesale price rigidity) and to incorporate the firm-driven benefits and costs from the manufacturer perspective. Sixth, whereas we are interested in past-price dependence, which occurs when retailers set prices on the basis of pricing history (whether that is everyday same pricing or alternating discounts), further research could perform a spectral analysis to decompose the nature and direction of autocorrelations in prices for different time frames and frequencies (e.g., Bronnenberg, Mela, and Boulding 2006; Lemmens, Croux, and Dekimpe 2007). Seventh, it would be fruitful to generalize the findings on the patterns of price adjustment to heterogeneous exogenous shocks (e.g., Levy, Dutta, and Bergen 2002) across products, retailers, and markets. Eighth, research in macroeconomics has built models that incorporate menu costs to explain aggregate patterns of price adjustment (e.g., Golosov and Lucas 2007). This research investigates the sensitivity of repricing rates to general inflation and transient monetary shocks. Golosov and Lucas (2007) assume that "under menu costs, any individual price will be constant most of the time and then occasionally jump to a new level" (p. 172), and subsequently, they remove the common sales promotion patterns from their data to obtain a good match between theory and empirics (p. 184). Moreover, they do not consider category- and brand-specific costs and benefits to price changes. Further research seems warranted to develop comprehensive models of price adjustment that account for these microeconomic characteristics. Finally, additional research could generalize our conceptual framework to different retail chains, industries, and countries and could further formalize the empirically demonstrated phenomena using analytical and structural models.

Technical Appendix A Calculating the Relative Influence of Retail-Price Drivers Using GFEVD

Nijs, Srinivasan, and Pauwels (2007) derive GFEVD estimates from VARX models. These models are well suited to measure the retail-pricing dynamics central to our study.

An 11-equation VARX model was estimated per category per store, with sales volume of the top three brands (S_i, i = 1, 2, 3), an other-brands composite (S_4), wholesale and retail prices of the top three brands (WP_i and RP_i, i = 1, 2, 3), and store traffic (ST, a proxy for interretailer competition) (Chintagunta 2002). In addition to the intercept (α), five sets of exogenous control variables are included in the model: (1) a deterministic trend t to capture the impact of omitted, gradually changing variables; (2) a set of dummy variables (HD) that equal one in the shopping periods around major holidays (Chevalier, Kashyap, and Rossi 2003); (3) four weekly dummy variables (SD) to account for seasonal fluctuations in sales or prices; (4) a step dummy variable for the impact of new product introductions (NP); and (5) feature (F) and display (D) variables for each brand (for a similar model setup, see Nijs et al. 2001; Pauwels, Hanssens, and Siddarth 2002; Srinivasan, Popkowski, and Bass 2000).[8] The VARX is specified in Table A1.

A stepwise procedure is used to determine the appropriate lag-length K, to eliminate redundant parameters, and to ensure that the model residuals are well behaved (for details, see Nijs, Srinivasan, and Pauwels 2007). The VARX parameters are then used to derive GFEVD estimates (Pesaran and Shin 1998). The GFEVD quantifies the dynamic influence of competitive retail prices, brand demand, wholesale price and competitive wholesale price, and category-management considerations on a brand's retail price. In essence, GFEVD provides a measure of the relative impact over time of shocks initiated by each of the individual endogenous variables in a VARX model, without the need for the researcher to specify a causal ordering among these variables (for a marketing application of FEVD, see Hanssens 1998). We derive the GFEVD estimates using the following equation:

$$\text{(A1)} \quad \theta_{ij}^{g}(n) = \frac{\sum_{l=0}^{n}\left[\psi_{ij}^{g}(l)\right]^{2}}{\sum_{l=0}^{n}\sum_{j=0}^{m}\left[\psi_{ij}^{g}(l)\right]^{2}},\ i, j = 1, \ldots, m,$$

where $\psi_{ij}^{g}(l)$ is the value of a generalized impulse response function following a one-unit shock to variable i on variable j at time l. By calculating GFEVD in this way, we ensure that the driver estimates are comparable across brands and categories (for details on the calculation, see, e.g., Dekimpe and Hanssens 1999).

---

[8] To avoid overparameterization, we include feature and display as exogenous variables (Pesaran and Smith 1998).

TABLE A1
VARX Equation

$$
\begin{bmatrix} S_{1,t} \\ S_{2,t} \\ S_{3,t} \\ S_{4,t} \\ RP_{1,t} \\ RP_{2,t} \\ RP_{3,t} \\ WP_{1,t} \\ WP_{2,t} \\ WP_{3,t} \\ ST_{t} \end{bmatrix}
=
\begin{bmatrix}
\alpha_{S1} + \delta_{S1}t + \sum_{h=1}^{H}\theta_{S1}^{h}HD_t + \sum_{s=2}^{S}\lambda_{S1}^{s}SD_t + \eta_{S1}NP_t \\
\alpha_{S2} + \delta_{S2}t + \sum_{h=1}^{H}\theta_{S1}^{h}HD_t + \sum_{s=2}^{S}\lambda_{S2}^{s}SD_t + \eta_{S2}NP_t \\
\alpha_{S3} + \delta_{S3}t + \sum_{h=1}^{H}\theta_{S3}^{h}HD_t + \sum_{s=2}^{S}\lambda_{S3}^{s}SD_t + \eta_{S3}NP_t \\
\alpha_{S4} + \delta_{S4}t + \sum_{h=1}^{H}\theta_{S4}^{h}HD_t + \sum_{s=2}^{S}\lambda_{S4}^{s}SD_t + \eta_{S4}NP_t \\
\alpha_{RP1} + \delta_{RP1}t + \sum_{h=1}^{H}\theta_{RP1}^{h}HD_t + \sum_{s=2}^{S}\lambda_{RP1}^{s}SD_t + \eta_{RP1}NP_t \\
\alpha_{RP2} + \delta_{RP2}t + \sum_{h=1}^{H}\theta_{RP2}^{h}HD_t + \sum_{s=2}^{S}\lambda_{RP2}^{s}SD_t + \eta_{RP2}NP_t \\
\alpha_{RP3} + \delta_{RP3}t + \sum_{h=1}^{H}\theta_{RP3}^{h}HD_t + \sum_{s=2}^{S}\lambda_{RP3}^{s}SD_t + \eta_{RP3}NP_t \\
\alpha_{CP1} + \delta_{CP1}t + \sum_{h=1}^{H}\theta_{CP1}^{h}HD_t + \sum_{s=2}^{S}\lambda_{CP1}^{s}SD_t + \eta_{CP1}NP_t \\
\alpha_{CP2} + \delta_{CP2}t + \sum_{h=1}^{H}\theta_{CP2}^{h}HD_t + \sum_{s=2}^{S}\lambda_{CP2}^{s}SD_t + \eta_{CP2}NP_t \\
\alpha_{CP3} + \delta_{CP3}t + \sum_{h=1}^{H}\theta_{CP3}^{h}HD_t + \sum_{s=2}^{S}\lambda_{CP3}^{s}SD_t + \eta_{CP3}NP_t \\
\alpha_{ST} + \delta_{ST}t + \sum_{h=1}^{H}\theta_{ST}^{h}HD_t + \sum_{s=2}^{S}\lambda_{ST}^{s}SD_t + \eta_{ST}NP_t
\end{bmatrix}
+ \sum_{k=1}^{K}
\begin{bmatrix}
\beta_{1,1}^{k} & \cdots & \beta_{1,10}^{k} \\
\beta_{2,1}^{k} & \cdots & \beta_{2,10}^{k} \\
\beta_{3,1}^{k} & \cdots & \beta_{3,10}^{k} \\
\beta_{4,1}^{k} & \cdots & \beta_{4,10}^{k} \\
\beta_{5,1}^{k} & \cdots & \beta_{5,10}^{k} \\
\beta_{6,1}^{k} & \cdots & \beta_{6,10}^{k} \\
\beta_{7,1}^{k} & \cdots & \beta_{7,10}^{k} \\
\beta_{8,1}^{k} & \cdots & \beta_{8,10}^{k} \\
\beta_{9,1}^{k} & \cdots & \beta_{9,10}^{k} \\
\beta_{10,1}^{k} & \cdots & \beta_{10,10}^{k} \\
\beta_{10,1}^{k} & \cdots & \beta_{10,10}^{k}
\end{bmatrix}
\begin{bmatrix} S_{1,t-k} \\ S_{2,t-k} \\ S_{3,t-k} \\ S_{4,t-k} \\ RP_{1,t-k} \\ RP_{2,t-k} \\ RP_{3,t-k} \\ WP_{1,t-k} \\ WP_{2,t-k} \\ WP_{3,t-k} \\ ST_{t-k} \end{bmatrix}
+ \sum_{l=0}^{L}
\begin{bmatrix}
\gamma_{1,1}^{l} & \cdots & \gamma_{1,6}^{l} \\
\gamma_{2,1}^{l} & \cdots & \gamma_{2,6}^{l} \\
\gamma_{3,1}^{l} & \cdots & \gamma_{3,6}^{l} \\
\gamma_{4,1}^{l} & \cdots & \gamma_{4,6}^{l} \\
\gamma_{5,1}^{l} & \cdots & \gamma_{5,6}^{l} \\
\gamma_{6,1}^{l} & \cdots & \gamma_{6,6}^{l} \\
\gamma_{7,1}^{l} & \cdots & \gamma_{7,6}^{l} \\
\gamma_{8,1}^{l} & \cdots & \gamma_{8,6}^{l} \\
\gamma_{9,1}^{l} & \cdots & \gamma_{9,6}^{l} \\
\gamma_{10,1}^{l} & \cdots & \gamma_{10,6}^{l} \\
\gamma_{11,1}^{l} & \cdots & \gamma_{11,6}^{l}
\end{bmatrix}
\begin{bmatrix} F_{1,t-l} \\ F_{2,t-l} \\ F_{3,t-l} \\ D_{1,t-l} \\ D_{2,t-l} \\ D_{3,t-l} \end{bmatrix}
+
\begin{bmatrix} \varepsilon_{S1,t} \\ \varepsilon_{S2,t} \\ \varepsilon_{S3,t} \\ \varepsilon_{S4,t} \\ \varepsilon_{RP1,t} \\ \varepsilon_{RP2,t} \\ \varepsilon_{RP3,t} \\ \varepsilon_{WP1,t} \\ \varepsilon_{WP2,t} \\ \varepsilon_{WP3,t} \\ \varepsilon_{ST,t} \end{bmatrix}
$$

Notes: Σ is the covariance matrix of the residuals $[\varepsilon_{S1,t}\ \varepsilon_{S2,t}\ \varepsilon_{S3,t}\ \varepsilon_{S4,t}\ \varepsilon_{RP1,t}\ \varepsilon_{RP2,t}\ \varepsilon_{RP3,t}\ \varepsilon_{WP1,t}\ \varepsilon_{WP2,t}\ \varepsilon_{2P3,t}\ \varepsilon_{ST,t}]'$.

The relative importance of the drivers is established on the basis of the GFEVD values at 26 weeks, which reduces sensitivity to short-term fluctuations. To evaluate the accuracy of our GFEVD estimates, we obtain standard errors using Monte Carlo simulations (see Benkwitz, Lütkepohl, and Wolters 2001).

Technical Appendix B
Bootstrap Algorithm to Correct the Standard Error Bias from OLS Estimation

To correct the parameter standard error bias we introduced when estimating Equations 1 and 2 with OLS, we use the following bootstrap algorithm:

Step 1: Select a sample of size n, with replacement, from the GFEVD estimates provided by Nijs, Srinivasan, and Pauwels (2007), where n is equal to the number of observations in the data set.

Step 2: Add measurement error based on Monte Carlo–simulated GFEVD estimates to each element of the sample. This step is repeated 250 times, each time creating a variation of the data set obtained in Step 1.

Step 3: Calculate parameter estimates θ^* for Equations 2 and A1 for each of the 250 augmented data sets created in Step 2.

We repeat Steps 1–3 250 times. The standard deviations across the 62,500 parameter vectors ($\theta^{*1}, \theta^{*2}, \ldots, \theta^{*62{,}500}$) are the unbiased standard errors for θ^{OLS} (for details, see Bradley and Tibshirani 1993).

REFERENCES

AMR Research Inc. (2000), "The Report on Retail Applications," (September), 1–20.

Basuroy S., Murali K. Mantrala, and Rockney G. Walters (2001), "The Impact of Category Management on Retailer Prices and Performance: Theory and Evidence," *Journal of Marketing*, 65 (October), 16–32.

Bell, David R., Jeongwen Chiang, and V. Padmanabhan (1999), "The Decomposition of Promotional Response: An Empirical Generalization," *Marketing Science*, 18 (4), 504–526.

Benkwitz, Alexander, Helmut Lütkepohl, and J. Wolters (2001), "Comparison of Bootstrap Confidence Intervals for Impulse Responses of German Monetary Systems," *Macroeconomic Dynamics*, 5 (1), 81–100.

Bergen, Mark, Mark Ritson, Shantanu Dutta, Daniel Levy, and Mark Zbaracki (2003), "Shattering the Myth of Costless Price Changes," *European Management Journal*, 21 (6), 663–69.

Bils, Mark and Peter J. Klenow (2004), "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112 (5), 947–85.

Blattberg, Robert C., G.D. Eppen, and J. Lieberman (1981), "A Theoretical and Empirical Evaluation of Price Deals for Consumer Nondurables," *Journal of Marketing*, 45 (Winter), 116–29.

——— and Scott A. Neslin (1990), *Sales Promotions: Concepts, Methods and Strategies*. Englewood Cliffs, NJ: Prentice Hall.

Blinder, Alan S. (1991), "Why Are Prices Sticky? Preliminary Results from an Interview Study (in Price Stickiness in Theory and Practice)," *American Economic Review*, 81 (2), 89–96.

———, E.R.D. Canetti, D.E. Lebow, and J.B. Rudd (1998), *Asking About Prices: A New Approach to Understanding Price Stickiness*. New York: Russell Sage Foundation.

Bolton, Ruth (1989), "The Relationship Between Market Characteristics and Promotional Price Elasticities," *Marketing Science*, 18 (2), 153–69.

Bradley, Efron and Robert J. Tibshirani (1993), *An Introduction to the Bootstrap: Monographs on Statistics and Applied Probability*, Vol. 57. New York: Chapman and Hall.

Bronnenberg, Bart J. and Vijay Mahajan (2001), "Unobserved Retailer Behavior in Multimarket Data: Joint Spatial Dependence in Market Shares and Promotion Variables," *Marketing Science*, 20 (3), 284–99.

———, Carl F Mela, and William Boulding (2006), "The Periodicity of Pricing," *Journal of Marketing Research*, 43 (August), 477–93.

BusinessWeek (2000), "The Power of Smart Pricing," (April 10), (accessed November 6, 2007), [available at http://www.businessweek.com/2000/00_15/b3676133.htm].

Carlton, Dennis W. (1986), "The Rigidity of Prices," *American Economic Review*, 76 (4), 637–58.

Chevalier, Judith A., Anil K. Kashyap, and Peter E. Rossi (2003), "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 93 (1), 15–37.

Chevalier, Michel and Ronald C. Curhan (1976), "Retail Promotions as a Function of Trade Promotions: A Descriptive Analysis," *Sloan Management Review*, 18 (1), 19–32.

Chintagunta, Pradeep K. (2002), "Investigating Category-Pricing Behavior at a Retail Chain," *Journal of Marketing Research*, 32 (May); 141–54.

Dekimpe, Marnik G. and Dominique M. Hanssens (1999), "Sustained Spending and Persistent Response: A New Look at Long-Term Marketing Profitability," *Journal of Marketing Research*, 36 (November), 397–412.

Dhar, Sanjay K., Stephen Hoch, and Nanda Kumar (2001), "Effective Category Management Depends on the Role of the Category," *Journal of Retailing*, 77 (2), 165–84.

Doornik, Jurgen A. (2002), *Ox: An Object-Oriented Matrix Language*, 4th ed. London: Timberlake Consultants.

Dutta, Shantanu, Mark Bergen, and Daniel Levy (2002), "Price Flexibility in Channels of Distribution: Evidence from Scanner Data," *Journal of Economic Dynamics and Control*, 26 (11), 1845–1900.

Fader, Peter S. and Leonard M. Lodish (1990), "A Cross-Category Analysis of Category Structure and Promotional Activity of Grocery Products," *Journal of Marketing*, 54 (July), 52–65.

Golosov, Mikhail and Robert E. Lucas (2007), "Menu Costs and Phillips Curve," *Journal of Political Economy*, 115 (2), 171–99.

Gordon, Robert E. (1981), "Output Fluctuations and Gradual Price Adjustment," *Journal of Economic Literature*, 19 (2), 493–530.

Greenleaf, Eric A. (1995), "The Impact of Reference Price Effects on the Profitability of Price Promotions," *Marketing Science*, 14 (Winter), 82–104.

Hanssens, Dominique M. (1998), "Order Forecasts, Retail Sales and the Marketing Mix for Consumer Durables," *Journal of Forecasting*, 17 (3–4), 327–46.

Information Resources Inc. (1997), *The Marketing Factbook Annual Report*. Chicago: Information Resources Inc.

Kahn, Barbara E. and Leigh McAlister (1997), *Grocery Revolution: The New Focus on the Consumer*. Reading, MA: Addison-Wesley.

Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1–22.

Kim, Sang Yong and Richard Staelin (1999), "Manufacturer Allowances and Retailer Pass-Through Rates in a Competitive Environment," *Marketing Science*, 18 (1), 59–76.

Kinberg, Yoram, Ambar G. Rao, and Melvin F. Shakun (1974), "A Mathematical Model for Price Promotions," *Management Science*, 20 (February), 948–55.

Kopalle, Praveen K., Ambar G. Rao, and João L. Assunção (1996), "Asymmetric Reference Price Effects and Dynamic Pricing Policies," *Marketing Science*, 15 (1), 60–85.

Krishna, Aradhna, Carl F. Mela, and Joel Urbany (2001), "Inertia in Pricing," working paper, College of Business Administration, University of Notre Dame.

Lach, Saul and Daniel Tsiddon (1996), "Staggering and Synchronization in Price Setting: Evidence from Multiproduct Firms," *American Economic Review*, 86 (5), 1175–96.

Lal, Rajiv (1990), "Price Promotions: Limiting Competitive Encroachment," *Marketing Science*, 9 (Summer), 247–62.

———, John D. Little, and Miguel Villas-Boas (1996), "A Theory of Forward Buying, Merchandising and Trade Deals," *Marketing Science*, 15 (1), 21–37.

Lemmens, Aurélie, Christophe Croux, and Marnik G. Dekimpe (2007), "Consumer Confidence in Europe: United in Diversity," *International Journal of Research in Marketing*, 24 (2), 113–27.

Levy, Daniel (2007), "Price Rigidity and Flexibility: New Empirical Evidence," *Managerial and Decision Economics*, 28 (6), 523–30.

———, Mark Bergen, Shantanu Dutta, and Robert Venable (1997), "The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains," *Quarterly Journal of Economics*, 112 (3), 791–825.

———, Shantanu Dutta, and Mark Bergen (2002), "Heterogeneity in Price Rigidity: Evidence from a Case Study Using Micro-Level Data," *Journal of Money Credit Banking*, 34 (1), 197–220.

———, ———, and Robert Venable (1998), "Price Adjustment at Multiproduct Retailers," *Managerial and Decision Economics*, 19 (2), 81–120.

Mankiw, N. Gregory (1985), "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly," *Quarterly Journal of Economics*, 100 (2), 529–37.

Mazumdar, Tridib and Purushottam Papatla (2000), "An Investigation of Reference Price Segments," *Journal of Marketing Research*, 37 (May), 246–58.

McAlister, Leigh (2008), "Cross-Brand Pass-Through: Fact or Artifact," *Marketing Science*, forthcoming

Midrigan, Virgiliu (2006), "Menu Costs, Multi-Product Firms and Aggregate Fluctuations," working paper, Department of Economics, Ohio State University.

Moorthy, Sridhar (2005), "A General Theory of Pass-Through in Channels with Category Management and Retail Competition," *Marketing Science*, 24 (1), 110–22.

Müller, George, Mark Bergen, Shantanu Dutta, and Daniel Levy (2006), "Private Label Price Rigidity During Holiday Periods," *Applied Economics Letters*, 13 (1), 57–62.

Murphy, Kevin M. and Robert H. Topel (1985), "Estimation and Inference in Two-Step Econometric Models," *Journal of Business Economic Statistics*, 3 (4), 370–79.

Narasimhan, Chakravarthi, Scott A. Neslin, and Subrata K. Sen (1996), "Promotional Elasticities and Category Characteristics," *Journal of Marketing*, 60 (April), 17–30.

Nijs, Vincent, Marnik G. Dekimpe, Jan-Benedict E.M. Steenkamp, and Dominique M. Hanssens (2001), "The Category Demand Effects of Price Promotions," *Marketing Science*, 20 (1), 1–22.

———, Shuba Srinivasan, and Koen Pauwels (2007), "Retail-Price Drivers and Retailer Profits," *Marketing Science*, 26 (4), 473–87.

Pauwels, Koen (2007), "How Retailer and Competitor Decisions Drive the Long-Term Effectiveness of Manufacturer Promotions for Fast-Moving Consumer Goods," *Journal of Retailing*, 83 (3), 297–308.

———, Dominique M. Hanssens, and S. Siddarth (2002), "The Long-Term Effects of Price Promotions on Category Incidence, Brand Choice, and Purchase Quantity," *Journal of Marketing Research*, 39 (November), 421–39.

——— and Shuba Srinivasan (2004), "Who Benefits from Store Brand Entry?" *Marketing Science*, 23 (3), 364–90.

Pesaran, Hashem H. and Yongcheol Shin (1998), "Generalized Impulse Response Analysis in Linear Multivariate Models," *Economic Letters*, 58 (1), 17–29.

——— and Ron P. Smith (1998), "Structural Analysis of Cointegrating VARs," *Journal of Economic Surveys*, 12 (5), 471–506.

Raju, Jagmohan S. (1992), "The Effect of Price Promotions on Variability in Product Category Sales," *Marketing Science*, 11 (3), 207–220.

———, V. Srinivasan, and Rajiv Lal (1990), "The Effects of Brand Loyalty on Competitive Price Promotional Strategies," *Management Science*, 36 (3), 276–304.

Rao, Ram (1991), "Pricing and Promotions in Asymmetric Duopolies," *Marketing Science*, 10 (Spring), 131–44.

Rotemberg, Julio J. (2002), "Customer Anger at Price Increases, Time Variation in the Frequency of Price Changes and Monetary Policy," NBER Working Paper No. 9320.

Slade, Margaret E. (1998), "Optimal Pricing with Costly Adjustment: Evidence from Retail-Grocery Prices," *Review of Economic Studies*, 65 (1), 87–107.

Srinivasan, Shuba, Koen Pauwels, Dominique Hanssens, and Marnik Dekimpe (2004), "Do Promotions Benefit Manufacturers, Retailers or Both?" *Management Science*, 50 (5), 617–29.

———, Peter T.L. Popkowski, and Frank M. Bass (2000), "Market Share Response and Competitive Interaction: The Impact of Temporary, Evolving and Structural Changes in Prices," *International Journal of Research in Marketing*, 17 (4), 281–305.

Varian, Hal R. (1980), "A Model of Sales," *American Economic Review*, 70 (4), 651–59.

Walters, Rockney G. (1989), "An Empirical Investigation into Retailer Response to Manufacturer Trade Promotions," *Journal of Retailing*, 65 (Summer), 253–72.

Wedel, Michel, Jie Zhang, and Fred Feinberg (2004), "A Model-Based Approach to Setting Optimal Retail Markups," working paper, Stephen M. Ross School of Business, University of Michigan.

Wolman, Alexander (2007), "The Frequency and Costs of Individual Price Adjustment," *Managerial and Decision Economics*, 28 (6), 531–52.

Zbaracki, Mark J., Mark Ritson, Daniel Levy, Shantanu Dutta, and Mark Bergen (2004), "Managerial and Customer Costs of Price Adjustment: Direct Evidence from Industrial Markets," *Review of Economics and Statistics*, 86 (2), 514–33.

Detelina Marinova, Jun Ye, & Jagdip Singh

Do Frontline Mechanisms Matter? Impact of Quality and Productivity Orientations on Unit Revenue, Efficiency, and Customer Satisfaction

This study identifies a frontline mechanism comprising autonomy, cohesion, and feedback that helps explain when and why the simultaneous pursuit of quality and productivity orientations has positive or negative effects on unit revenue, efficiency, and customer satisfaction. An empirical test of the proposed framework using data from 423 employees in 30 strategic business units and longitudinal unit-level performance data indicates that frontline autonomy mediates the positive impact of productivity and quality orientations on unit revenue and customer satisfaction and their negative impact on unit efficiency. Feedback amplifies the influence of frontline autonomy by simultaneously enhancing its positive effect on satisfaction and its negative effect on efficiency. In contrast, unit cohesion strengthens the positive effect of frontline autonomy on revenue and customer satisfaction without augmenting its negative effect on unit efficiency. The results urge managers to shift their focus toward unit-level mechanisms to find clues for managing strategic dilemmas that stem from multiple goal pursuit in face-to-face service settings.

Keywords: frontline mechanism, quality orientation, productivity orientation, revenue, efficiency, customer satisfaction

Attaining bottom-line results that are both superior and sustainable becomes more difficult when an organization faces conflict between two or more desirable business orientations (Anderson, Fornell, and Rust 1997; Bateson 1985; Grönroos and Ojasalo 2004; Rust, Moorman, and Dickson 2002). This is especially true for service organizations. Consider the simultaneous adoption of a quality orientation, which focuses on achieving a high level of service quality, and a productivity orientation, which focuses on attaining high internal efficiency. Empirical evidence suggests that service companies that adopt both orientations have conflicting strategic goals of efficient resource deployment and effective customer retention (Bateson 1985). Anderson, Fornell, and Rust (1997) report that service companies experience a negative correlation between productivity and satisfaction, whereas product companies do not (–.30 versus .15), and the simultaneous pursuit of productivity and satisfaction also has a negative, albeit marginal, effect on profitability (see Bucklin 1978). Grönroos and Ojasalo (2004, p. 415) characterize the cumulative evidence as indicative of a service productivity dilemma, such that "increased internal efficiency following the introduction of more cost effective and seemingly more productive processes" may lead to lower service quality and lost revenues.

The practitioner literature is replete with examples of service organizations struggling with this dilemma. For example, Starbucks's chairman, Howard Schultz, recently took the unusual step of publicly releasing a "blunt" memo to his executives, warning that Starbucks's drive for growth and efficiency (e.g., automatic espresso machines instead of in-store coffee grinding) was compromising the quality of the customer experience, the very reason for the firm's phenomenal success and ability to charge premium prices (Adamy 2007). Although this is an unusual step for Starbucks's chairman, the underlying dilemma is not unusual for many service organizations (Hindo 2007; Singh 2000).

However, practitioners and researchers agree that despite trade-offs in their simultaneous adoption, both pro-

---

Detelina Marinova is Assistant Professor of Marketing, Department of Marketing, Robert J. Trulaske, Sr., College of Business, University of Missouri (e-mail: marinovad@missouri.edu). Jun Ye is Assistant Professor of Marketing, Lundquist College of Business, University of Oregon (e-mail: junye@uoregon.edu). Jagdip Singh is Professor of Marketing, Weatherhead School of Management, Case Western Reserve University (e-mail: jxs16@po.cwru.edu). The authors contributed equally to the development of this article. The basic work on this article was completed as part of Jun Ye's doctoral studies at Case Western Reserve University. The support of National Science Foundation Grant No. SES-0080567, awarded to Professors Jagdip Singh and Gil Preuss, with Sister Nancy Linenkugel, is also acknowledged.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (March 2008), 28–45

ductivity and quality are necessary orientations, especially in services with a significant labor component, such as banking, education, health care, and entertainment (Mittal et al. 2005). Consider the case of health care. For years, the industry focused on quality without regard for productivity; now, the market is demanding that it increases productivity without sacrificing quality. Health care providers, regulators, and insurance providers are publicizing report cards that evaluate and rank hospital on both cost (lower costs are associated with higher productivity) and quality criteria. The goal is "to steer consumers to better-performing hospitals—and reward hospitals and doctors for providing higher-quality, cost-effective care" (Landro 2004, p. D1). Parallel market pressures are evident in other services as well (Ante and Sager 2002; Carey and Frangos 2005).

Similarly, researchers have argued that the productivity–quality trade-off should not be taken to imply that service firms should not seek improvements in both productivity and customer satisfaction (Anderson, Fornell, and Rust 1997; Grönroos and Ojasalo 2004). Various solutions have been conceptually proposed to address the productivity–quality dilemma, including process improvement through information technology applications (Anderson, Fornell, and Rust 1997), better management of frontline human capital through autonomy (Singh 2000), and continuous learning in the frontline units (Grönroos and Ojasalo 2004). Recent empirical evidence suggests that firms that manage to achieve both productivity and quality orientations enjoy superior long-term financial returns, highlighting that successful implementation remains a key challenge (Mittal et al. 2005). Collectively, these studies suggest that service organizations have a strategic imperative to pursue both improved productivity and improved quality. Understanding the mechanisms that intervene between these orientations and performance appears essential to manage the apparent conflict that arises in adopting productivity and quality orientations.

The purpose of this study is to investigate when and why the simultaneous pursuit of productivity and quality goals results in positive or negative outcomes for a service business unit by identifying an intervening mechanism between a strategic business unit's (SBU's) strategic orientations and its performance outcomes. Building on Day and Wensley's (1988) notion of dynamic capabilities as warehoused in a service unit's practices and routines, we conceptualize an intervening mechanism as including autonomy, feedback, and cohesion among a unit's frontline employees. Such mechanisms, referred to as self-management practices, have been shown to explain unit outcomes, such as service quality, satisfaction, and customer loyalty (Batt 1999; De Jong, De Ruyter, and Lemmink 2004).

We focus on frontline employees because their direct contact with customers makes their performance a key factor in a service unit's effectiveness (Heskett, Sasser, and Schlesinger 1997). To our knowledge, no study to date has examined the mediating role of frontline mechanisms on unit performance and satisfaction outcomes. Although prior research has often made the organization the unit of analysis, we follow the lead of strategy work and recent services research in choosing the SBU as the focal unit of analysis (De Jong, De Ruyter, and Lemmink 2004). Given the heterogeneity in routines and practices in a diversified, multiunit organization, the SBU is a more suitable level for analysis because it mitigates aggregation biases. This is especially relevant when the SBUs are functionally diverse to serve a range of customer needs, such as in health care (e.g., radiology, surgery, ob–gyn) and financial services (e.g., personal banking, insurance; Bucklin 1978). We test the proposed framework using survey data from 423 employees in 30 SBUs, as well as longitudinal unit-level archival and customer satisfaction data for the corresponding units.

Conceptual Model and Hypotheses

Three key elements of our proposed model (see Figure 1) are notable. First, longitudinal archival data provide information on an SBU's unit revenue, efficiency, and customer satisfaction, which allows for the possibility of contrasting effects and trade-offs among these outcomes. Unit revenue refers to an SBU's total revenue per unit of output, unit efficiency refers to the total labor cost per unit of output, and satisfaction refers to customers' evaluations of delivered service.

Second, the proposed frontline mechanism is indicated by the interrelationship of autonomy, cohesion, and feedback for a unit's frontline employees. We posit that this mechanism mediates the effects of productivity and quality orientations on service outcomes.

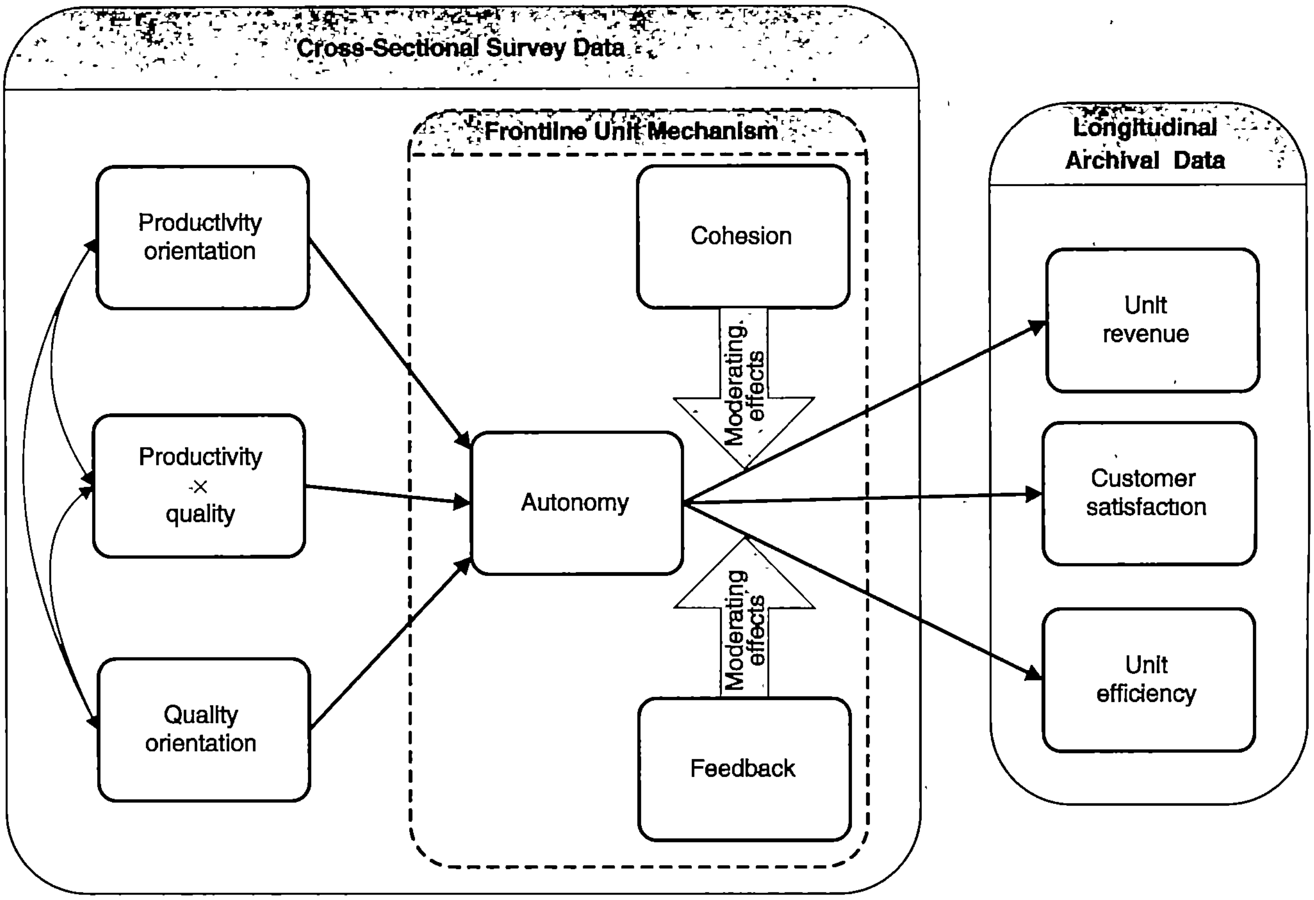
Third, drawing on control theory (Eisenhardt 1985; House and Rizzo 1972) and supporting empirical evidence (Bandura and Cervone 1986; Baron and Kreps 1999; Durham and Bartol 2000; Larkin and Larkin 1996), we propose that management practices are key manifestations of a unit's strategic orientations. Management practices are actions and directives that flow from management priorities and guidelines for regulating employee efforts toward desired goals (Hambrick and Mason 1984; Simons 2000).

Productivity and Quality Orientations

We define productivity orientation as management practices oriented toward maintaining high levels of internal efficiency and cost control, and we define quality orientation as management practices oriented toward achieving high levels of service quality and customer satisfaction. Prior studies indicate that within an organizational unit, employees' interpretation of managerial decisions and priorities may result in perceptions of practices in use that differ from what managers intended (Gregory 1983; Lytle, Hom, and Mokwa 1998; Schneider 1987). Employees make sense of managerial practices through an iterative process of receiving managerial inputs, responding to managerial expectations, and adjusting their actions on the basis of managerial feedback (Schneider et al. 2003). Because employee attitudes and performance are direct precursors to unit outcomes (Harter, Schmidt, and Hayes 2002; Heskett, Sasser, and Schlesinger 1997), researchers have argued that employee perceptions are useful for the study of organizational orientations and their performance impact. Lytle,

UNIVERSITY
OF
CALCUTTA

FIGURE 1
The Hypothesized Role of Frontline Mechanisms on the Impact of Productivity and Quality Orientations on Unit Revenue, Efficiency, and Customer Satisfaction



Hom, and Mokwa (1998, p. 458) go so far as to claim that an "orientation of an organization is what its employees perceive it to be." Thus, we conceptualize that productivity and quality orientations are reflected in managerial practices as reported by the frontline employees of an organizational unit.

The Role of Frontline Mechanisms Involving Unit Autonomy, Cohesion, and Feedback

Most studies have examined the direct effect of strategic orientations on performance (Gatignon and Xuereb 1997; Rust, Moorman, and Dickson 2002; Voss and Voss 2000). To our knowledge, however, no study has yet examined or identified the mechanisms mediated by frontline employees that facilitate or hinder the successful adoption of a unit's productivity and quality orientations (Mittal et al. 2005). We argue that without well-aligned frontlines, a unit's strategic orientations are less likely to generate the anticipated outcomes. We develop hypotheses for these frontline mechanisms in shaping an SBU's productivity–quality trade-offs. We do not specifically hypothesize direct effects of unit practices on outcomes, because such hypotheses represent no incremental insights to those in prior research. However, we include the direct-effect paths in our estimations. To examine the incremental contribution of the proposed model, we compare it with a direct-effects model that omits the intervening frontline mechanisms and represents extant research in the field.

The frontline autonomy mechanism. According to job design theory, the structure and design of a job directly regulates employees' effort to achieve organizational goals (Cooper and Foster 1971; Cummings 1978). Typically, managers have a choice between external control (through hierarchical supervision, work scheduling, and standardization) and internal control (through the promotion of self-regulation and self-determination among frontline employees) (Langfred and Moye 2004; Mills and Ungson 2003). Managers attempt to choose the control system that will be most effective in goal attainment, given the specifics of job requirements.

We reason that by altering the level of autonomy afforded in frontline jobs, managers can vary the control mechanism along a continuum from external control (low autonomy) to internal control (high autonomy). In self-determination theory, autonomy is a central variable for supporting an internal mechanism that enhances frontline employee motivation and participation (Deci and Ryan 1987; Hackman and Oldham 1976). Designing jobs with greater autonomy shifts the locus of control from the manager (external) to the employee (internal).

Consider the managerial choice of control when quality goals are emphasized. In this case, managers are likely to require that employees provide customized or individualized solutions to meet heterogeneous and unpredictable customer needs. This requirement entails designing frontline jobs for higher levels of autonomy. This is because managers are unable to anticipate the variety of possible customer needs and must rely on employees' "local knowledge" to provide solutions that address those needs. In other words, the appropriate control mechanism for maintaining a high level of service quality is delegation of decision-making power and accountability (Argote 1982; Bowen and Jones 1986; Zeithaml and Bitner 1996).

In contrast, when managers emphasize a productivity orientation, they prioritize productivity goals, such as increasing the output–input ratio and controlling costs. In this case, managers are likely to try to limit variability in frontline performance and to support standardized solutions for customers' varying needs (Bowen and Lawler 1992). Hierarchical supervision and limitations on autonomous action by frontline employees—in other words, external control mechanisms—are likely to be preferred managerial choices because they are effective at reducing variability in productivity goal attainment, resulting in a negative influence on frontline autonomy.

In turn, frontline autonomy is hypothesized to be positively associated with a unit's revenue and customer satisfaction (Anderson, Fornell, and Rust 1997; Mittal et al. 2005; Rust, Moorman, and Dickson 2002) and negatively associated with a unit's efficiency (Levy and Pashler 2001; Osman and Moore 1993). Frontline autonomy promotes effective response to quality challenges, customer needs, and recovering service failures (Bowen and Lawler 1992), resulting in increased customer satisfaction and revenues. However, autonomy can engender unnecessary variability, slow the speed of service delivery, and increase employees' cognitive burden for task selection and strategy (Langfred and Moye 2004; Pashler 1998; Rubinstein, Meyer, and Evans 2001), thus hurting unit efficiency.

On the basis of the preceding insights, we propose that autonomy is likely to mediate positively the effect of a unit's quality orientation on outcomes related to service quality (e.g., customer satisfaction and revenue) but mediate negatively the effect of productivity orientation on efficiency-related outcomes. Thus:

> H_1: Frontline autonomy positively mediates the influence of quality orientation on customer satisfaction and revenue generation such that (a) the greater the unit's quality orientation, the higher is the level of perceived autonomy among its frontline employees and (b) perceived autonomy is positively associated with customer satisfaction and unit revenue.
>
> H_2: Frontline autonomy negatively mediates the influence of productivity orientation on unit efficiency such that (a) the greater the unit's productivity orientation, the lower is the level of perceived autonomy among its frontline employees and (b) perceived autonomy is negatively associated with unit efficiency.

Furthermore, different organizational units may vary in the degree to which they emphasize productivity- and quality-oriented practices. Unit managers may not fully recognize productivity–quality trade-offs; they may believe that they need to emphasize both because of top management directives, or they may believe that they can effectively overcome productivity–quality conflicts. Other managers may place greater emphasis on one orientation while keeping both in focus because it fits the needs of their unit or provides an effective way of diffusing the tension between the two orientations. Regardless of the reasons, a simultaneous pursuit of multiple orientations is conceptualized in most previous studies by including a product term involving the focal orientations, as we show in Figure 1 (e.g., Rust, Moorman, and Dickson 2002). The notion of trade-offs is captured by the dysfunctional influence of the product term, whereas the individual orientations on their own may have a positive, functional effect (Anderson, Fornell, and Rust 1997).

We posit that when quality and productivity orientations are simultaneously emphasized, unit managers need to grant frontline employees increased autonomy to help them cope with the inherent tension of quality- and productivity-oriented tasks (Anderson, Fornell, and Rust 1997; Singh 2000). That is, the simultaneous adoption of productivity and quality orientations entails multiple comparison standards and task interrelatedness, which require greater discretion and coping efforts on the part of employees. For example, frontline employees may be required to reconcile service quality feedback (e.g., not enough customer attention) that conflicts with productivity feedback (e.g., service delivery must be accomplished more quickly). In such instances, managers are likely to give frontline employees greater autonomy so that the employees can effectively reconcile this conflict through creative problem solving. Thus, we expect that the interaction effect of productivity and quality orientations on frontline autonomy will be positive. Moreover, when combined with our prediction for linear effects (H_1 and H_2), we expect contrasting patterns for the impact of productivity and quality orientations. Specifically, we expect that the negative linear association between productivity orientation and autonomy will become less negative or even positive when managers simultaneously implement a quality orientation. In contrast, the positive association between quality orientation and autonomy will be enhanced with an increased, simultaneous emphasis on productivity orientation.

> H_3: The interaction between a unit's quality and productivity orientations is positively associated with autonomy among a unit's frontline employees.

The autonomy–performance moderators. Designing jobs for frontline autonomy risks loss of some organizational control, even when the job demands require it. Mills and Ungson (2003, p. 146) point out that control loss is associated with attrition of goal-directed outcomes because employees "may not understand what they are supposed to do or may not choose to do what is expected of them." Managers attempt to manage this control loss by supporting other mechanisms that moderate the autonomy–performance relationship. Moderators of interest involve unit characteristics that bolster the expected positive effect of frontline autonomy on unit outcomes, thus curtailing control losses. We investigate two frontline unit characteristics—unit cohesion and performance feedback—that are posited to moderate the autonomy–SBU outcomes relationship. Specifically, we argue that performance feedback operates by directing frontline employees' behaviors toward what they are supposed to do and that unit cohesion helps by maintaining frontline employees' commitment to these goals. Moreover, the pursuit of multiple, usually conflicting goals in frontline context adds another layer of complexity to the moderating effect of performance feedback and unit cohesion.

The moderating effect of unit cohesion. We posit that in face-to-face service settings, unit cohesion amplifies both autonomy's positive effect on revenue and customer satisfaction and its negative effect on efficiency. To understand this effect, consider the mechanism by which unit cohesion influences group processes and functions in the pursuit of multiple goals. As a recent meta-analysis demonstrates (Mullen and Copper 1994), a positive relationship between group cohesion and performance occurs because of group members' commitment to shared goals and their ability to self-regulate their collective behavior to that end (see Stogdill 1972; Terborg, Castore, and DeNinno 1976). However, if the group's shared goals are inconsistent with organizational goals, a more cohesive group will perform relatively poorly compared with a less cohesive group, especially when the group has autonomy in goal pursuit and regulation (see Cherrington et al. 2001).

Thus, the effects of unit cohesion depend on how well the employees' shared goals match the organization's goals. In face-to-face service settings, several studies have confirmed that frontline workers and their organizations are aligned in their pursuit of quality goals (e.g., providing high levels of medical care) but differ in their priority for productivity goals (e.g., providing cost-efficient care; Begat, Ellefsen, and Severinsson 2005; Summer and Townsend-Rocchiccioli 2003; Weinberg 2003).[1] That is, whereas management tends to place high priority on productivity goals, frontline employees tend to focus on high quality of service and believe that the former comes at the price of the latter. These beliefs result in discrepancy between the importance of productivity goals for the organization and its frontline employees. Subsequently, we confirm this empirically with a follow-up study (see the Appendix).[2] This discrepancy is sometimes reflected in unusual business practices. Consider the case of the common airline practice of overbooking flights to increase efficiency. In a recent article, Baily (2007) reported that gate agents (frontline employees) who were appalled by such overbooking practices took the matter into their own hands by creating phantom reservations (e.g., using names such as the airline chief executive officer) to keep a flight from being oversold. Situated at the interface with the customer, gate agents were motivated to minimize dealing with bumped, dissatisfied passengers and to keep the customer experience positive.

Accordingly, we reason that cohesion in frontline units is likely to enhance the functional effect of autonomy on unit revenue and customer satisfaction because frontline units and management are aligned in their quality goals and emphasis on service quality. In contrast, the discrepancy between the importance management and employees place on productivity goals means that unit cohesion is likely to amplify the negative effect of autonomy on unit efficiency. Recall that we argued that unit autonomy reduces efficiency because it diverts cognitive resources to task selection and strategy. With increased unit cohesion, we posit that frontline employees are more likely to direct their cognitive resources toward aligned goals (quality) and away from discrepant goals (efficiency). Thus:

> H_4: Units with greater cohesion among frontline employees also demonstrate a stronger (a) positive effect of frontline autonomy on unit revenue, (b) positive effect of frontline autonomy on customer satisfaction, and (c) negative effect of frontline autonomy on unit efficiency.

The moderating effect of performance feedback. We posit parallel hypotheses for the moderating effect of feedback on the relationship between autonomy and unit outcomes. We define feedback as the information managers provide to unit employees to direct employees' service performance. Our logic is based on studies of feedback when multiple goals are being pursued. A widely shared assertion is that feedback improves outcomes in high-involvement systems because feedback identifies performance gaps (e.g., when and where performance is below goal levels), and autonomous employees can regulate their efforts to reduce gaps (Kernan and Lord 1990; Klein 1989). However, according to feedback intervention theory (Kluger and DeNisi 1996), people assign higher priority to feedback gaps related to goals that are more important and involve higher commitment (Locke and Latham 1990; Terborg and Miller 1978), and this hierarchical priority directs people's effort and resource deployment. Thus, employees are less likely to allocate extra effort and resources to reduce performance gaps for goals they deem to be less important or attractive and more likely to ignore or reject managerial feedback related to those goals (Hollenbeck and Williams

---

[1] There is evidence that service occupations self-select employees who tend to be intrinsically motivated for service effectiveness. For example, empirical studies show that when organizations adopt a customer-oriented culture, frontline employees show high job satisfaction and commitment (Ruekert 1992; Siguaw, Brown, and Widing 1994).

[2] In the follow-up study, we collected data from unit managers. By matching manager and employee data, we computed a D statistic (Cronbach and Gleser 1953) as an indicator of goal incongruence between frontline employees and managers for each unit and separately for productivity and quality orientations. The results indicate that managers and frontline employees differ significantly in the importance of productivity goals (mean D = 2.03) compared with quality goals (mean D = .50), with a significant discrepancy for the former (t = 13.12, d.f. = 46, $p < .01$).

1987; Taylor, Fisher, and Ilgen 1984). This implies that autonomous frontline employees in face-to-face service settings are likely to give less credence to managerial feedback related to productivity goals and, instead, focus efforts on reducing performance gaps related to quality goals. This is because they tend to place a higher priority on quality goals than productivity goals, and they are aligned with managers with regard to quality but not productivity goals. Thus, increasing performance feedback is likely to direct the effort of autonomous frontline employees toward quality goals and away from productivity goals, which suggests a negative autonomy–efficiency relationship under this condition.

> H_5: Units whose frontline employees receive more performance feedback demonstrate a stronger (a) positive effect of frontline autonomy on unit revenue, (b) positive effect of frontline autonomy on customer satisfaction, and (c) negative effect of frontline autonomy on unit efficiency.

Method

Research Design

Setting. We chose nonprofit hospitals as our research setting for several reasons. First, the health care industry operates in a dynamic and competitive environment that demands both high quality and productivity (Fennell and Alexander 1993). Traditionally focused exclusively on their service mission, nonprofit hospitals have begun emphasizing productivity in addition to quality (Aiken, Sochalski, and Lake 1997). Therefore, the questions posed in this research are especially relevant for them. Second, hospitals are organized around distinct units that perform frontline functions by interacting directly with patients but that differ in terms of specific tasks (e.g., the tasks of a surgery unit differ from those of a pediatric unit) as well as composition and skill base of their employees (e.g., radiologists have different skills than obstetric nurses). Thus, hospital units can be viewed as relatively independent SBUs and treated as the units of analysis. We selected one hospital in the northeastern United States for this study. Selection of one hospital with multiple frontline units affords variation in the focal variables, while minimizing potentially confounding variables due to differences in organizational contexts.

Data sources and sampling. Data were collected from multiple sources. We used key informant data for the proposed independent, mediating, and moderating variables; longitudinal archival data for unit service performance, including revenue, and efficiency; and longitudinal satisfaction data collected from patients. Because frontline mechanisms can be best understood by securing data from frontline employees who are directly involved in these practices, every employee with direct patient interactions was considered a possible key informant and was mailed a questionnaire packet that included (1) a letter from the researchers describing the purpose of the study, (2) a seven-page questionnaire, (3) a return postage-paid envelope, and (4) a lottery-card-based incentive (for a demographic profile, see Table 1). We obtained complete responses from 729 employees in 47 units, for a response rate of 45%. A comparison of early and late respondents revealed no significant mean differences for study variables, with the exception of cohesion ($F = 3.99$, $p = .046$), which reduces to nonsignificance after we control for the number of comparisons and overall Type I error rate.

We obtained archival data on unit quarterly service performance for a three-year period; frontline employee survey data collection occurred at the midpoint; that is, the archival data covered a year-and-a-half window before and after the survey data collection. This ensured that we captured the downstream effect of unit practice orientations on performance. We systematically removed data for units that were not independent cost centers, reorganized or merged with other units, did not have meaningful revenue data, or did not provide complete data on all required measures for computing unit revenue and efficiency. After matching the individual frontline employee and unit-level data, we had complete observations for 423 frontline employees in 30 units, which we retained for the unit revenue and efficiency

TABLE 1
Demographic Profile of Employees from Participating Service Organization (All Numbers Are Percentages)

Age (In Years)	%	Education	%
<25	6.0	High school	2.2
25–35	32.9	Technical certification	9.0
36–45	27.3	Associate degree	43.8
46–55	24.9	Some college	8.3
56–65	7.9	College	30.7
>65	1.0	Graduate school	6.0

Tenure (In Years)	%	Salary (In $/Year)	%
<2	16.7	<10,000	6.4
2–5	23.6	10,000–29,999	29.8
6–10	12.7	30,000–49,999	50.2
11–15	15.8	50,000–69,999	13.4
16–20	8.8	70,000–89,999	.2
>20	22.4		

analysis. We obtained customer satisfaction data through a marketing research company that continuously measures customer satisfaction through exit surveys. Complete customer satisfaction data were available for the customers of 10 units with 193 frontline employee observations.

Measurement and Operationalization

Productivity and quality orientations. Employees provided key informant data on management's productivity and quality orientations. We used four items to measure productivity-oriented management practices (PROD) in the unit: (1) Management decisions reflect serious intentions to improve hospital productivity (mgmt1), (2) management urges employees to cut hospital costs (mgmt2), (3) employees are expected to focus on increasing efficiency (mgmt3), and (4) enhancing organizational productivity is a priority for the management (mgmt4). Scale validation testing resulted in the dropping of one item (mgmt1). Likewise, we measured quality-oriented management practices (QUAL) with four items: (1) Management places the highest priority on delivering the best-quality care (mgmt5), (2) management views medical errors as opportunities to improve the quality of medical care (mgmt6), (3) management focuses on ensuring the highest levels of patient satisfaction (mgmt7), and (4) management views patient complaints as opportunities to improve future patient satisfaction (mgmt8). We used a five-point Likert scale that ranged from "strongly disagree" to "strongly agree." The reliabilities of the productivity and quality orientations were .83 and .88, respectively.

Constructs comprising the frontline unit mechanism. We measured unit autonomy (AUT) with a three-item, five-point Likert scale. We derived this measure from Hackman and Oldham's (1976) job diagnostic survey scale, which has received wide attention in the marketing literature (e.g., Singh 1998). Employees indicated their agreement or disagreement regarding whether they had (1) the freedom to do their job in the way they thought best (auto1), (2) opportunities to do whatever was needed to provide quality patient care (auto2), and (3) activities that allowed for independent thought and action (auto3).[3] We also assessed unit cohesion (COH) with a three-item, five-point Likert scale that was based on the group cohesiveness scale that Beehr (1976) developed. Employees indicated their agreement or disagreement with the following statements: (1) Employees have a shared sense of community and purpose (coh1), (2) one can count on assistance from fellow employees (coh2), and (3) one can trust coworkers to lend a hand in need (coh3). Subsequent analysis resulted in the dropping of one item (coh1). For performance feedback (FEED), employees rated the following three items using a five-point Likert scale that ranged from "never" to "very frequently": (1) Managers provide useful feedback to individual employees (feed1), (2) managers discuss methods for improving individual performance (feed2), and (3) managers provide data on individual performance (feed3). The reliabilities for the autonomy, cohesion, and feedback scales were .87, .94, and .96; respectively.

Unit service performance. We used two measures of unit performance—revenue (REVE) and cost efficiency (COST)—that were based on longitudinal unit-level quarterly archival data provided by each hospital for the period January 2000–December 2002. We computed REVE as follows:

$$(1)\quad \text{REVE}_t = \text{Unit gross revenue}_t/\text{Unit equivalent patient days}_t,$$

where t indexes the revenue for any given quarter and REVE estimates the revenue generated from one equivalent patient day. We computed equivalent patient day for each unit by estimating the following proportion: unit total gross patient revenue/hospital gross revenue per equivalent patient day. In turn, the hospital gross revenue per equivalent patient day is as follows: [total gross inpatient revenue/(number of inpatients × length of average inpatient stay)]. Hospitals routinely use the unit equivalent-patient-day measure to assess performance across units (both inpatient and outpatient units) and to determine unit-specific budgets.

In addition, we computed a measure of unit efficiency as a ratio of the cost of human factors in service delivery to an equivalent patient day. In service organizations, the human capital cost is one of the most significant expenses, and efficiency gains in service delivery flow from incremental reductions in the cost of human factors per equivalent patient day. For interpretability, we reversed the sign of this ratio to obtain a positively increasing scale for efficiency. Similarly, adjusting the cost of human factors by equivalent patient days for a given unit accounts for variations in units' size, labor intensity, nature of services, patient mix, and other unit-specific factors. Thus, we define unit efficiency as follows:

$$(2)\quad \text{EFFI}_t = (-\text{Total labor cost}_t/\text{Unit equivalent patient days}_t).$$

Although the unit revenue and efficiency measures provide an adjustment for differences in unit characteristics, we performed additional analysis to rule out the potentially confounding effects of unobservable variables. Specifically, beginning with the adjusted quarterly time-series data for REVE_t and EFFI_t for each unit, we used econometric procedures to extract the variability arising from autocorrelation and seasonality effects. To control for the effects of unobservable variables, we modeled a first-order autocorrelation (e.g., Boulding and Staelin 1995; Jacobson 1990) and included a fourth-difference term for seasonality effects. Thus, we estimated the following time-series cross-sectional models for each unit j:

$$(3)\quad \text{REVE}_t = \beta_0 + \beta_1\text{REVE}_{t-1} + \beta_2\text{REVE}_{t-4} + \varepsilon_t, \text{ and}$$

$$(4)\quad \text{EFFI}_t = \lambda_0 + \lambda_1\text{EFFI}_{t-1} + \lambda_2\text{EFFI}_{t-4} + \xi_t,$$

where t represents time (range: 1–12) and REVE and EFFI represent unit revenue and efficiency, respectively, as we calculated previously.[4] Consistent with prior research

---

[3] We analyzed the data after omitting auto2 because of a reviewer's concern about potential confounding with quality orientation. The pattern of results remained the same, as we show subsequently.

[4] As an additional analysis and check, we modified Equations 3 and 4 to capture only "downstream" data by including a dummy

(Bayus, Erickson, and Jacobson 2003), we retained β_0 and λ_0 as the corrected estimates of unit revenue and efficiency for use in testing the hypotheses, as follows:

$$\text{REVE}_i = \beta_0, \text{ and} \tag{5}$$

$$\text{EFFI}_i = \lambda_0. \tag{6}$$

We derived customer satisfaction (CS) for each of the ten units from the unit-level quarterly satisfaction data for the same three-year period as the financial data. We used patients' responses to the question "Overall, how would you rate the care you received at the unit?" to compute the customer satisfaction score. The patients responded using a five-point Likert scale that ranged from "poor" to "excellent." We derived the customer satisfaction index by the following procedures: For each quarter t, we computed a mean score Satisfaction_t based on the patients' rating on the overall care they received in the unit:

$$\text{Satisfaction}_t = \frac{1 \times X_{1t} + 2 \times X_{2t} + 3 \times X_{3t} + 4 \times X_{4t} + 5 \times X_{5t}}{X_{1t} + X_{2t} + X_{3t} + X_{4t} + X_{5t}}, \tag{7}$$

where t represents time (quarter) and ranges from 1 to 12; 1, 2, 3, 4, and 5 represent the categories "poor," "fair," "good," "very good," and "excellent," respectively; and X_{1t}, X_{2t}, X_{3t}, X_{4t}, and X_{5t} represent the frequency of the category indicated in the subscript at time t, respectively. Finally, we derived the customer satisfaction indicator (CS) for each unit by averaging the satisfaction rating across the 12 quarters:

$$\text{CS} = \frac{1}{12}\sum_{t=1}^{12} \text{Satisfaction}_t. \tag{8}$$

Method of Analysis

The analytical approach involved measurement assessment of the key constructs and testing the hypothesized model. Table 2 summarizes the key descriptive statistics for the studied constructs.

Measurement analysis. We used standard confirmatory factor analysis procedures to assess the psychometric properties of the key constructs at the unit level. We explicitly focused on the evidence for the convergent and discriminant validity of the study constructs.

Hypothesized model analysis. To test the hypotheses, we estimated the proposed model and a rival model that omitted the unit mechanisms. The rival model was a direct-effects model in which unit strategic orientations are posited to influence service performance directly, in accordance with prior research. To account for the multilevel structure of the data (frontline employees nested within units) and model individual unobserved heterogeneity, we used a random-parameters model (Greene 2007), which allows for between- and within-unit effects. Thus, this method enables us to capture the effects of the hypothesized variables (i.e., frontline autonomy, unit cohesion, and feedback) on performance outcomes by modeling parameter heterogeneity.

$$\begin{aligned}\text{REVE}_{ij} = {} & \delta_{0j} + \delta_{1j}\text{Aut}_{ij} + \delta_{2j}\text{Aut}_{ij} \times \text{Coh}_{ij} \\ & + \delta_{3j}\text{Aut}_{ij} \times \text{Feed}_{ij} + \delta_{4j}\text{Coh}_{ij} + \delta_{5j}\text{Feed}_{ij} + \delta_{6j}\text{Qual}_{ij} \\ & + \delta_{7j}\text{Prod}_{ij} + \delta_{8j}\text{Prod}_{ij} \times \text{Qual}_{ij} + \varepsilon_{ij},\end{aligned} \tag{9a}$$

$$\delta_{nj} = \varphi_n + u_{nj}; \tag{9b}$$

$$\begin{aligned}\text{EFFI}_{ij} = {} & \gamma_{0j} + \gamma_{1j}\text{Aut}_{ij} + \gamma_{2j}\text{Aut}_{ij} \times \text{Coh}_{ij} + \gamma_{3j}\text{Aut}_{ij} \times \text{Feed}_{ij} \\ & + \gamma_{4j}\text{Coh}_{ij} + \gamma_{5j}\text{Feed}_{ij} + \gamma_{6j}\text{Qual}_{ij} + \gamma_{7j}\text{Prod}_{ij} \\ & + \gamma_{8j}\text{Prod}_{ij} \times \text{Qual}_{ij} + \varsigma_{ij},\end{aligned} \tag{10a}$$

$$\gamma_{nj} = \omega_n + u_{nj}; \tag{10b}$$

$$\begin{aligned}\text{SAT}_{ij} = {} & \alpha_{0j} + \alpha_{1j}\text{Aut}_{ij} + \alpha_{2j}\text{Aut}_{ij} \times \text{Coh}_{ij} \\ & + \alpha_{3j}\text{Aut}_{ij} \times \text{Feed}_{ij} + \alpha_{4j}\text{Coh}_{ij} + \alpha_{5j}\text{Feed}_{ij} \\ & + \alpha_{6j}\text{Qual}_{ij} + \alpha_{7j}\text{Prod}_{ij} + \alpha_{8j}\text{Prod}_{ij} \times \text{Qual}_{ij} + \nu_{ij},\end{aligned} \tag{11a}$$

$$\alpha_{nj} = \phi_n + u_{nj}; \text{ and} \tag{11b}$$

TABLE 2
Descriptive Statistics and Intercorrelations for the Study Constructs

	1	2	3	4	5	6	7	8	M	SD
1. Quality orientation	.83								3.60	.75
2. Productivity orientation	.16	.88							3.80	.66
3. Unit autonomy	.22**	–.07	.87						3.70	.81
4. Unit cohesion	.17	–.02	.24**	.94					3.87	.97
5. Performance feedback	.33**	.02	.31*	.16	.96				3.01	.50
6. Unit revenue (REVE)[a]	.11	.08	.17	.08	.16	1.00			.02	2.05
7. Unit efficiency (EFFI)[a]	.12	.03	.17*	.07	.18	.18	1.00		.15	1.50
8. Customer satisfaction (SAT)[a]	–.02	.07	–.13	–.07	–.14	.49	–.46	1.00	4.08	.21

*$p < .05$.
**$p < .01$.
[a] Unit revenue, unit efficiency, and customer satisfaction measures are based on actual financial archival data and thus are assumed to have a reliability of 1.
Notes: We estimated all correlations at the unit level after aggregating the individual-level data. Customer satisfaction statistics are based on a subset of units (10) and employees (193)

---

variable that takes the value of 0 if t = 1–6 or the value of 1 if t = 7–12. If the coefficient of the dummy variable was significant, we retained it as the appropriate performance index. Using this performance index produced no significant difference in estimated coefficients. We thank an anonymous reviewer for suggesting this.

$$\text{(12a)} \quad \text{AUT}_{ij} = \theta_{0j} + \theta_{1j}\text{Qual}_{ij} + \theta_{2j}\text{Prod}_{ij} + \theta_{3j}\text{Prod}_{ij} \times \text{Qual}_{ij} + \rho_{ij},$$

$$\text{(12b)} \quad \theta_{nj} = \tau_n + u_{nj},$$

where i and j denote individual and unit, respectively; AUT = autonomy, COH = cohesion, FEED = feedback, QUAL = quality orientation, and PROD = productivity orientation; the coefficients δ_{nj}, γ_{nj}, α_{nj}, and θ_{nj} capture individual-specific unobserved heterogeneity within units; and u_{nj} are ~ $N(0, \sigma^2)$ and denote unit-specific variances. Note that the coefficients δ_{nj}, γ_{nj}, α_{nj}, and θ_{nj} can change randomly between units because of unobserved within-unit heterogeneity. In turn, φ_n, ω_n, ϕ_n, and τ_n capture the between-units effects estimated in Equations 9b, 10b, 11b, and 12b, which account for unit-specific variances (u_{nj}).

The rival direct-effects model involved the estimation of the following equations:

$$\text{(13a)} \quad \text{REVE}_{ij} = \beta_{0j} + \beta_{1j}\text{Qual}_{ij} + \beta_{2j}\text{Prod}_{ij} + \beta_{3j}\text{Prod}_{ij} \times \text{Qual}_{ij} + \zeta_{ij},$$

$$\text{(13b)} \quad \beta_{nj} = \psi_n + u_{nj};$$

$$\text{(14a)} \quad \text{EFFI}_{ij} = \upsilon_{0j} + \upsilon_{1j}\text{Qual}_{ij} + \upsilon_{2j}\text{Prod}_{ij} + \upsilon_{3j}\text{Prod}_{ij} \times \text{Qual}_{ij} + o_{ij},$$

$$\text{(14b)} \quad \upsilon_{nj} = \sigma_n + u_{nj}; \text{ and}$$

$$\text{(15a)} \quad \text{SAT}_{ij} = \pi_{0j} + \pi_{1j}\text{Qual}_{ij} + \pi_{2j}\text{Prod}_{ij} + \pi_{3j}\text{Prod}_{ij} \times \text{Qual}_{ij} + \iota_{ij},$$

$$\text{(15b)} \quad \pi_{nj} = \eta_n + u_{nj}.$$

In estimating the hypothesized and rival models, we were sensitive to multicollinearity issues. Consequently, we used an instrumental variable for quality orientation orthogonal to the other independent variables in the analysis. We also considered the possibility that heterogeneity in employee characteristics influenced the modeled relationships. For example, employees with higher education and work experience tend to be in positions with higher autonomy than those with lower education and work experience. To rule out such alternative explanations, we included several control variables (i.e., income, education level, and work experience) based on prior research that identified these characteristics as influential in group processes and performance outcomes (Campion, Medsker, and Higgs 1993; Gladstein 1984). To the extent that we obtained empirical support for the focal hypotheses after modeling unobserved individual-specific heterogeneity as well as the discussed control variables, our theory is robust to potential alternative explanations.

Results

Measurement Analysis

A confirmatory factor analysis of the study constructs with individual-level data and using the EQS software yielded the following fit statistics (see Table 3): $\chi^2 = 774.35$, d.f. = 105, $p = .11$; normed fit index = .88; nonnormed fit index = .97; comparative fit index = .98; standardized root mean square residual = .10; and root mean square error of approximation = .07 (90% confidence interval = .00–.11). On statistical, absolute, and relative fit, as well as substantive grounds, the posited measurement model provides a good fit to the data. Table 3 provides further support for the convergent and discriminant validity of the constructs. For example, the estimated loadings for the relationship between individual indicants and their underlying construct are, without exception, large and significant (t-values > 3.5, $p < .01$). In addition, the reliability estimates are large and significant, ranging from .83 to .95, with an average reliability index of .89, which exceeds the conventional .70 criterion. In terms of discriminant validity, the epistemic correlations among the study constructs for individual-level data range from .01 to .81, with none approaching unity. Consistent with this, the 95% confidence intervals for construct correlations do not include unity. In accord with Fornell and Larcker's (1981) criterion for discriminant validity, the variance extracted exceeds the average variance shared for each study construct. Together, the preceding evidence provides robust support for the convergent and discriminant validity of study constructs.

Hypothesized Model Analysis

Table 4 shows the results of model fit tests and coefficient estimates for the hypothesized and direct-effects (rival) models obtained using the Limdep software. Compared with the null model, the hypothesized model provides a significant improvement in fit for unit revenue ($\chi_{11\text{d.f.}} = 41.19$, $p < .01$), efficiency ($\chi_{11\text{d.f.}} = 37.31$, $p < .01$), and satisfaction ($\chi_{11\text{d.f.}} = 35.94$, $p < .01$), and the rival model is significant for unit revenue ($\chi_{6\text{d.f.}} = 14.53$, $p < .05$) but nonsignificant for efficiency ($\chi_{6\text{d.f.}} = 9.12$, $p > .10$) and customer satisfaction ($\chi_{6\text{d.f.}} = 10.6$, $p > .10$). Furthermore, a nested-model comparison of the hypothesized and direct-effects (rival) models based on the likelihood ratio test indicates that the former provides a significantly superior fit to the data compared with the latter for unit revenue ($\chi_{5\text{d.f.}} = 80.78$, $p < .01$), unit efficiency ($\chi_{5\text{d.f.}} = 33.54$, $p < .01$), and customer satisfaction ($\chi_{5\text{d.f.}} = 84$, $p < .01$). This is also confirmed by the lower Akaike information criterion value in the hypothesized model than the rival model in all three cases. We conclude that frontline autonomy, cohesion, and feedback constitute a regulating mechanism for frontline units and contribute significantly to the model fit and to an explanation of service performance and customer satisfaction.

The mediating role of frontline unit autonomy. In support of H_{1a} and H_{2a}, the results indicate that quality orientation has a significant, positive effect ($\theta_1 = .22$, $p < .01$) and that productivity orientation has a significant, negative effect ($\theta_2 = -.08$, $p < .01$) on perceived autonomy among frontline employees. Furthermore, the quality × productivity interaction term has a significant, positive effect on frontline autonomy ($\theta_3 = .12$, $p < .01$), in support of H_3. Finally, in support of H_{1b} and H_{2b}, autonomy has a direct

TABLE 3
Factor Loadings and Measurement Properties of Various Study Constructs

	Loading	t-Value	Reliability	Variance Extracted	Average R^2
Productivity Orientation			.83	.62	.31
Mgmt2	.99	7.63			
Mgmt3	.56	3.77			
Mgmt4	.76	5.32			
Quality Orientation			.88	.65	36
Mgmt5	.75	5.22			
Mgmt6	.73	5.02			
Mgmt7	.93	7.13			
Mgmt8	.81	5.78			
Unit Autonomy			.87	.70	.20
Auto1	.92	6.70			
Auto2	.84	5.92			
Auto3	.74	5.08			
Unit Cohesion			.94	.88	.05
Coh2	.90	6.70			
Coh3	.98	7.26			
Performance Feedback			.95	.88	.28
Feed1	.90	6.88			
Feed2	.99	8.07			
Feed3	.91	6.98			

Notes: Loading = standardized coefficient estimate based on the maximum likelihood estimation method using EQS software. t-values greater than 1.96 indicate significant effects at $p = .05$ for a two-tailed test. Composite reliability and variance extracted are based on Fornell and Larcker's (1981) formula. Average R^2 is the average variance shared between this construct and all other constructs; it is computed as the mean of squared correlations.

positive effect on unit revenue ($\delta_1 = .19$, $p < .05$) and customer satisfaction ($\alpha_1 = .52$, $p < .05$) and a negative effect on unit efficiency ($\gamma_1 = -.10$, $p < .10$).[5]

The moderating role of frontline unit cohesion and feedback. Table 4 indicates that unit cohesion has a significant, positive moderating effect on the relationship between frontline autonomy and unit revenue ($\delta_2 = .20$, $p < .05$) and that between autonomy and customer satisfaction ($\alpha_2 = .35$, $p < .05$). These results support H_{4a} and H_{4b}. However, unit cohesion does not have a significant moderating effect on the relationship between frontline autonomy and unit efficiency ($\gamma_2 = .04$, n.s.), so H_{4c} is not supported. In support of H_{5b} and H_{5c}, feedback has a significant, positive moderating effect on the relationship between frontline autonomy and customer satisfaction ($\alpha_3 = .34$, $p < .05$) and a significant, negative moderating impact on the relationship between autonomy and efficiency ($\gamma_3 = -.18$, $p > .05$). However, feedback does not have a significant moderating effect on the relationship between frontline autonomy and unit revenue ($\delta_3 = -.07$, $p > .10$); thus, H_{5a} is not supported.

Direct impact of quality and productivity orientations. To test the validity of the proposed unit mechanism, we also included the direct effects of quality and productivity orientations. Table 4 indicates that productivity orientation has a significant, positive effect on unit revenue ($\beta_2 = .27$, $p < .01$) but not on unit efficiency ($\upsilon_2 = -.06$, $p > .10$) or customer satisfaction ($\pi_2 = -.13$, $p > .10$). Furthermore, quality orientation has no direct effect on unit revenue ($\beta_1 = .12$, $p > .10$), unit efficiency ($\upsilon_1 = .03$, $p > .10$), or customer satisfaction ($\pi_1 = .27$, $p > .10$). The quality × productivity interaction term is similarly nonsignificant for unit revenue ($\beta_3 = -.04$, $p > .10$), efficiency ($\upsilon_3 = -.17$, $p > .10$), and customer satisfaction ($\pi_3 = -.06$, $p > .10$).

Discussion

This study identifies a frontline mechanism that mediates the effect of strategic orientation on an SBU's service outcomes and is instrumental in explaining when and why the simultaneous pursuit of productivity and quality goals results in positive or negative service outcomes. The proposed frontline mechanism, in which frontline autonomy acts as the key mediator and unit cohesion and feedback are salient moderators, is found to be crucial in transmitting the impact of quality and productivity orientations and blunting the productivity–quality trade-offs that result from the pursuit of multiple goals. Thus, our study sheds new light on the link between strategic orientations and service outcomes and introduces a potential mechanism for managing the productivity–quality tension and improving the effectiveness of service companies. Before we discuss these results further, we identify several limitations of our research.

---

[5] We also empirically tested for moderation effects. We estimated the influence of interaction terms involving autonomy and quality orientation (effect on unit revenue = .09, t = 1.13; effect on efficiency = –.02, t = –.19; both *ps* > .10) as well as autonomy and productivity orientation (effect on unit revenue = .08, t = 1.06; effect on efficiency = .05, t = .77; both *ps* > .10). None of the moderation effects were significant.

TABLE 4
Results from Model Comparisons and Parameter Estimation

	Dependent Variables													
	Direct-Effects Model						Hypothesized Model							
	Unit Revenue		Unit Efficiency		Customer Satisfaction		Autonomy		Unit Revenue		Unit Efficiency		Customer Satisfaction	
Independent Variables	B	t	B	t	B	t	B	t	B	t	B	t	B	t
Quality orientation	.12	1.29	.03	.38	.27	1.69	.22***	5.50	–.06	–.72	.06	.97	.16	1.48
Productivity orientation	.27**	2.37	–.06	–.65	–.13	–.73	–.08***	–3.32	.01	.01	.09	1.35	–.15	–1.32
Productivity × quality	–.04	1.57	.17	–.52	–.06	.24	.12***	4.70	.14	1.70	–.25***	–4.53	–.35**	–2.08
Unit autonomy									.19*	1.91	–.10*	–1.83	.52**	2.08
Unit cohesion									.18*	1.83	.06	.91	.01	.06
Performance feedback									–.32***	–3.48	–.11**	–2.05	.14	.74
Autonomy × cohesion									.20**	2.23	.04	.67	.35*	1.92
Autonomy × feedback									–.07	–.73	–.18**	–2.22	.34*	1.93
Income	.05	.68	.08	1.12	.04	.53	–.04	–1.25	.21**	2.70	.09	1.59	–.05	–.80
Years of education	–.03***	–3.67	–.01	–.21	.07	1.95	–.01	–.18	–.13	–1.41	.06	.49	.04	.28
Work experience	.10	1.51	–.00	–.37	.13	1.41	–.02	–.77	–.13	–1.44	.06	.38	.03	.11
N_i, N_j	423, 30		423, 30		193, 10		423, 30		423, 30		423, 30		193, 10	
Log-likelihood	–751.36		–782.26		–579.18		–1074.71		–710.97		–765.49		–537.18	
Likelihood ratio test (d.f.)	14.53 (6)		9.12 (6)		10.6 (6)		68.20 (6)		41.19 (11)		37.31 (11)		35.94 (11)	
Akaike information criterion	2.44		.82		2.10		.38		1.09		.60		1.82	

*Significant at the .1 level.
**Significant at the .05 level.
***Significant at the .01 level.

First, the small sample size for the unit revenue and efficiency models has the potential to reduce statistical power and inflate Type II errors. To address this issue, we conducted analysis on longitudinal performance within units to derive appropriate performance indexes that increase the ability to detect a signal in the subsequent structural estimation of the model. Overall, we obtained statistical support for seven of the nine hypotheses, which alleviates concerns of inadequate power. Nevertheless, the small sample size warrants future replication and validation studies. Furthermore, customer satisfaction data were available only for 10 of the 30 SBUs. Empirically, we employ 193 data points for estimation, and the results indicate that the model explains a significant amount of the variation in customer satisfaction and provides statistical support for all hypotheses. Substantively, we suggest caution in generalizing the findings beyond the study setting at this time. Balancing the demands of multilevel data collection (involving individual frontline employee and unit-level data) with the study objectives, we believe that insights obtained are reasonable as an initial step in understanding the proposed frontline mechanism. Further testing with larger samples and different settings would be instrumental in generalizing the results. Second, we suggest caution against generalizing our findings beyond health care service organizations at this time. Although we expect that a comparable pattern of effects will emerge in other service contexts, validation in different settings is needed. Third, we recognize that other unmeasured variables may be involved in the proposed frontline unit mechanism. We reason that to the extent such variables are correlated with the control variables in our analysis and/or uncorrelated with the focal variables, our results are likely robust. In addition, we controlled for individual unobserved heterogeneity, which alleviates concerns of bias due to omitted interacting variables. Finally, the inclusion of autocorrelation specification in our estimations addresses the contemporaneous correlation problem in business performance models (Jacobson 1990). Nevertheless, our proposed model invites inclusion of other theoretically grounded mechanisms to map the generalizability and to extend the applicability of our findings.

Limited Direct Effects of Quality and Productivity Orientations on Service Performance

Prior research has focused mainly on the direct effects of strategic orientations on such unit outcomes as financial performance and has found supporting evidence (Gatignon and Xuereb 1997; Rust, Moorman, and Dickson 2002; Voss and Voss 2000). Our findings clarify and define the limits of this body of work by demonstrating that the direct effects are likely to be limited after we account for a frontline mediation mechanism, which is omitted in much prior research. Specifically, our results show that both productivity and quality orientations have no significant linear direct effects on unit revenue, unit efficiency, or customer satisfaction in face-to-face service settings. Furthermore, the interaction of productivity and quality has a significant, negative direct effect on both unit efficiency ($\gamma_3 = -.25, p < .01$) and customer satisfaction ($\alpha_1 = -.35, p < .01$). That is, unit efficiency and customer satisfaction decrease with increasing simultaneous emphasis on productivity and quality goals. This trade-off effect is consistent with prior research (Oliva and Sterman 2001; Rust, Moorman, and Dickson 2002), which shows service performance attrition under dual productivity–quality orientations.

Importantly, we demonstrate that the negative direct effect of the interaction between quality and productivity orientations on customer satisfaction is neutralized by a positive indirect effect mediated by the frontline mechanism. Specifically, the total marginal effect of quality orientation on satisfaction, given a high level of productivity orientation (+2 SD), is $\partial SAT/\partial Qual = -.47 + .16Coh + .16Feed$. Thus, depending on the level of frontline unit cohesion and performance feedback, the effect of quality orientation on customer satisfaction ranges from –1.11 for low cohesion and low feedback (both –2 SD) to +.17 for high cohesion and high feedback (both +2 SD). Furthermore, the effect of quality orientation on unit revenue becomes positive as a result of managing the simultaneous emphasis on quality and productivity orientation through the proposed frontline mechanism. The total marginal effect of quality orientation on unit revenue, given high levels of productivity orientation, is captured by $\partial REV/\partial Qual = .08 + .08Coh$ and ranges from –.08 for low cohesion to +.24 for high cohesion.

These findings have several implications. First, they suggest that the results of studies that focus on the simple direct effects of strategic orientations on SBU outcomes are likely to be incomplete, if not misleading, as long as the indirect and mediated effects are omitted. Second, the results indicate that consideration of a mediating frontline mechanism enhances understanding of the influence of productivity and quality orientations. Third, they suggest that whether the simultaneous adoption of productivity and quality orientations increases unit revenue and customer satisfaction depends less on the intensity with which the orientations are adopted and more on the amplifying and mitigating effects of the frontline mediation mechanism in face-to-face service settings. For example, increasing frontline autonomy can mitigate the negative effect of the simultaneous adoption of productivity and quality orientation on unit revenue. Likewise, together, cohesion and feedback can alter the direction and magnitude of the effects of strategic orientations on important SBU outcomes. Thus, studies that omit frontline mechanisms are likely to produce biased results driven by a contextually unspecified aggregation of mediated and unmediated effects, which will limit their usefulness for managers who want to increase productivity and quality in face-to-face service settings.

Influence of Frontline Mechanisms Based on Autonomy, Cohesion, and Feedback

Our results support the theory that frontline mechanisms play a crucial role in linking an SBU's strategic orientations and service outcomes. Specifically, unit autonomy acts as a key mediator in translating productivity and quality orientations into unit revenue, unit efficiency, and customer satisfaction. Consistent with our expectations, we found that though a quality orientation is positively associated with frontline autonomy, a productivity orientation is negatively associated with it. Furthermore, a simultaneous emphasis on quality and productivity has an incremental positive association with frontline autonomy. Finally, frontline autonomy has a significant, direct positive effect on unit revenue and customer satisfaction but a significant, negative effect on unit efficiency.

Several implications follow from these results. A greater emphasis on a quality orientation or a dual (quality and productivity) orientation is associated with increased frontline unit autonomy, whereas a greater emphasis on a productivity orientation is associated with lower levels of frontline autonomy. This is consistent with the notion that managers attempt to match the autonomy afforded to the front lines with the service demands of the job. The more autonomous units tend to achieve higher revenue but lower efficiency than units with lower levels of autonomy.

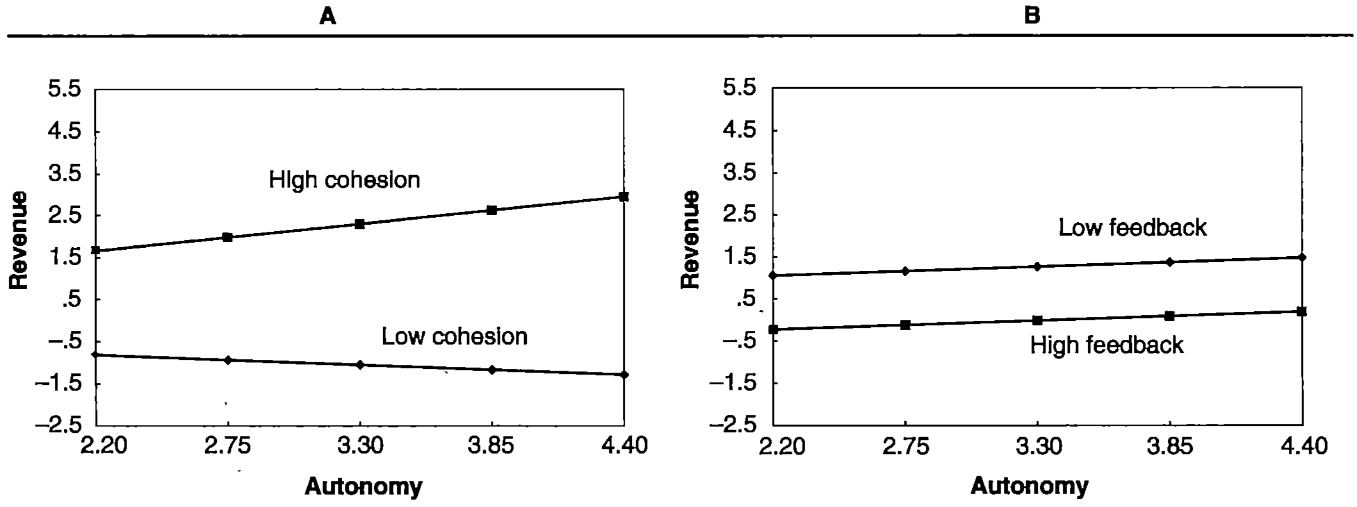
Our results pinpoint the locus of productivity–quality trade-offs in frontline autonomy. When guided by a quality orientation, units design frontline work for increased autonomy to gain unit revenue advantages, but in the process, they incur efficiency losses. In contrast, when guided by a productivity orientation, units design frontline work to limit autonomy, which results in increased efficiency but also significant revenue loss. A dual emphasis on both quality and productivity fails to eliminate trade-offs but shifts them toward higher revenue and customer satisfaction and lower efficiency. Thus, by providing evidence that quality and productivity orientations are predictably related to autonomy and that autonomy is linked to unit revenue, efficiency, and customer satisfaction, our results highlight the theoretical and pragmatic significance of frontline mediating mechanisms in face-to-face service settings.

In addition, our findings suggest that unit cohesion has a significant moderating effect on the frontline autonomy–performance link. We expected that unit cohesion would increase commitment to shared goals and unit members' tendency to self-regulate their collective behavior. Our results show that unit cohesion amplifies the positive impact of frontline autonomy on both unit revenue (slope changes from –.21 to .59) and customer satisfaction (slope changes from –.18 to 1.22; see Panel A of Figure 2 and Figure 3, respectively).

However, unit cohesion does not simultaneously increase the negative impact of autonomy on unit efficiency, which we also expected. This suggests that building unit cohesion in the front lines can help partly mitigate the inherent trade-offs in pursuing both quality and productivity orientation because it can enhance the positive mediation effects of autonomy on unit revenue and customer satisfaction without amplifying the negative mediation effect on unit efficiency. Importantly, maximal benefits from unit cohesion can be obtained only in the presence of frontline autonomy. Overall, these results suggest that the unit cohesion mechanism is important for building organizational capabilities that allow for the successful pursuit of multiple orientations.

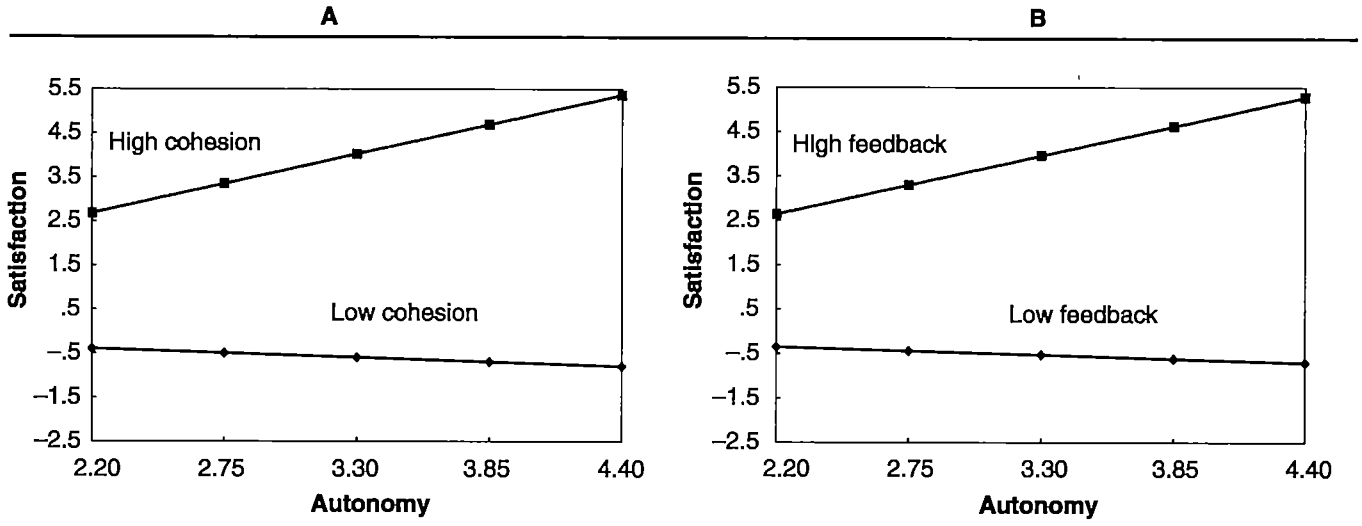
Finally, we expected that performance feedback would moderate the autonomy–performance relationship and found mixed evidence. Feedback amplifies frontline autonomy's positive effect on customer satisfaction (slope changes from –.16 to 1.20) and its negative effect on unit efficiency (slope changes from –.26 to –.46; see Panel B of Figure 3 and Figure 4, respectively). However, feedback has no significant moderating effect on the autonomy–unit revenue relationship (see Panel B of Figure 2). Thus, the interaction between frontline autonomy and performance feedback accentuates the customer satisfaction–efficiency trade-off. We reasoned that the influence of performance

FIGURE 2
Moderating Effects of Unit Cohesion and Performance Feedback on Frontline Autonomy → Revenue



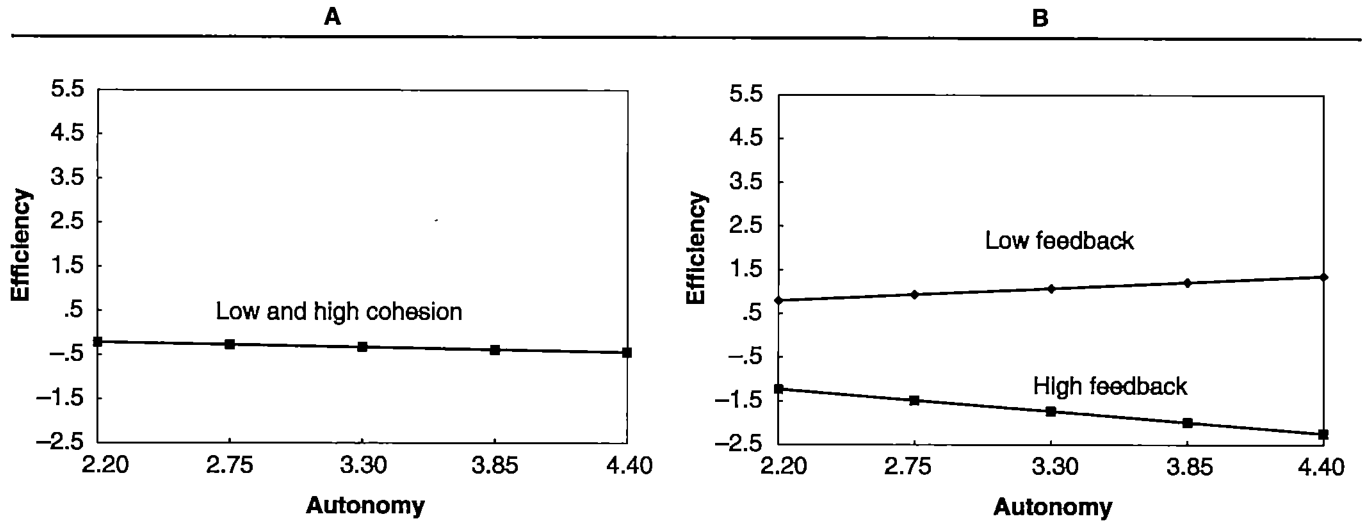
Notes: Panel A: The total effect of frontline autonomy on unit revenue is enhanced by the high versus low cohesion (±2 SD). The equations plotted are as follows: Unit revenue = –.36 – .21 × frontline autonomy (low unit cohesion), and unit revenue = .36 + .59 × frontline autonomy (high unit cohesion). Panel B: The total effect of frontline autonomy on unit revenue results in a smaller intercept for high versus low feedback (±2 SD) but equivalent slopes. The equations plotted are as follows: Unit revenue = 64 + .19 × frontline autonomy (low feedback), and unit revenue = –.64 + .19 × frontline autonomy (high feedback).

FIGURE 3
Moderating Effects of Unit Cohesion and Performance Feedback on Frontline Autonomy → Customer Satisfaction



Notes: Panel A: The total effect of frontline autonomy on satisfaction is enhanced by high versus low unit cohesion (±2 SD). The equations plotted are as follows: Customer satisfaction = –.18 × frontline autonomy (low unit cohesion), and customer satisfaction = 1.22 × frontline autonomy (high unit cohesion). Panel B. The total effect of frontline autonomy on satisfaction is enhanced by high versus low feedback (±2 SD). The equations plotted are as follows: Customer satisfaction = –.16 × frontline autonomy (low feedback), and customer satisfaction = 1.20 × frontline autonomy (high feedback).

FIGURE 4
Moderating Effects of Unit Cohesion and Performance Feedback on Frontline Autonomy → Efficiency



Notes: Panel A: The total effect of frontline autonomy on unit revenue remains the same across the level of unit cohesion (±2 SD). The equations plotted are as follows: Unit efficiency = –.10 × frontline autonomy (low unit cohesion), and unit efficiency = –.10 × frontline autonomy (high unit cohesion). Panel B: The total effect of frontline autonomy on unit efficiency is enhanced by low versus high feedback (±2 SD). The equations plotted are as follows: Unit efficiency = .22 + .26 × frontline autonomy (low feedback), and unit efficiency = –.22 – .46 × frontline autonomy (high feedback).

feedback in the presence of autonomy would likely vary with the frontline employees' shared alignment of multiple goals. In a face-to-face service setting, frontline employees and their organizations are aligned in the pursuit of quality goals, but they differ in their priority for productivity goals (e.g., Weinberg 2003). Consistent with this, performance feedback enhanced the functional effect of the aligned goal (e.g., quality through autonomy) and the dysfunctional effect of the misaligned goal (e.g., productivity through autonomy). In the absence of frontline autonomy, feedback has a negative effect on both unit revenue and unit efficiency in face-to-face service settings.

Managerial Implications and Concluding Notes

For managers who want or need to pursue strategic goals of productivity and quality simultaneously, our study offers some useful insights into managing the strategic dilemmas that stem from multiple goal pursuit in face-to-face service settings. First, while providing support for previous research that suggests inherent trade-offs in the simultaneous pursuit of productivity and quality goals (Anderson, Fornell, and Rust 1997; Rust, Moorman, and Dickson 2002; Singh 2000), our study shows that a consideration of unit mechanisms is necessary to understand fully the scope of these trade-offs. The magnitude of productivity–quality trade-offs does not simply depend on the intensity with which managers pursue these strategic orientations. Rather, the trade-offs are significantly amplified or mitigated by unit-level frontline mechanisms involving autonomy, cohesion, and feedback. For example, a clear and compelling finding of our study is that, together, frontline autonomy, cohesion, and feedback can bolster the influence of quality orientation by a factor equivalent to 1.25 standard deviations for customer satisfaction (e.g., from –1.11 to +.17) and more than .30 standard deviations for unit revenue (e.g., from –.08 to +.24). Based on current revenue data, a change of .30 standard deviations corresponds to incremental yearly revenue of more than $20 million for an SBU with 20 patients on the average per day.[6] Likewise, a change of 1.25 standard deviations in customer satisfaction scores can make the difference between acceptable and excellent service performance. The influence of frontline mechanisms is significant.

Second, although managers risk productivity–quality trade-offs when they design frontline jobs for autonomy, unit cohesion and performance feedback work to modulate the nature and degree of such trade-offs. For example, increased unit cohesion potentially eliminates the productivity–quality trade-off effect for unit revenue (e.g., from –.08 to .24) and substantially diminishes the trade-off for customer satisfaction (e.g., from –.79 to –.32), while leaving the effect on unit efficiency unaltered. However, although increased feedback weakens the impact of the productivity–quality trade-off on customer satisfaction (e.g., from –.79 to –.15), it amplifies the trade-off effect for unit efficiency (e.g., from –.38 to –.70).[7] Overall, consideration of both unit cohesion and performance feedback factors is needed to neutralize the trade-off effects on unit revenue and customer satisfaction. It is our assessment that attending to such unit-level mechanisms is critical for understanding and managing productivity–quality trade-offs in service organizations.

Third, readers may take our findings regarding the negative influence of autonomy on efficiency as evidence of inescapable trade-offs. This is a reasonable interpretation that undoubtedly presents a rather pessimistic picture of the strategic dilemmas inherent in simultaneous pursuit of productivity and quality goals. As a counterpoint, however, managers may ask whether it is strategically imperative to manage for unit revenue and customer satisfaction, for unit efficiency, or both. To the extent that customer satisfaction represents a superordinate goal, efficiency based on human capital costs may be important, but it is likely to be a red herring if pursued too aggressively. Nevertheless, our results reveal a revenue–efficiency tension that is different from the productivity–quality trade-off and that warrants additional research.

Fourth, our results urge managers and researchers to shift focus away from strategic orientations and toward unit-level mechanisms to find clues for managing strategic dilemmas, such as productivity–quality trade-offs. Modern business organizations face dilemmas such as the struggle between efficiency and creativity in their pursuit of competitive advantage in various market settings (Hindo 2007). In this sense, our research is in step with a growing body of work that highlights the significance of strategy-to-performance or "thinking–doing" gaps (Mankins and Steele 2005; Noble and Mokwa 1999). Explaining such gaps may be the key to resolving paradoxes and building a sustainable competitive advantage. We identified a frontline mechanism comprising autonomy, cohesion, and feedback that we view as a starting point for such research. An avenue for further research is the process of learning in the front lines and harnessing the resultant unique experiential knowledge at the customer interface to create innovative strategies and subsequent unit-level practices. We hope that further research uses this foundation to investigate other mediating variables that illuminate the pathways that link strategic productivity and quality orientations to unit revenue, customer satisfaction, efficiency, and other desired business outcomes.

Appendix
Using the D Statistic to Assess Employee–Manager Goal Incongruence

Unit managers and employees were asked to rate from 1 ("less important") to 5 ("extremely important") three quality-related goal items—(1) highest quality of patient care, (2) highest level of customer satisfaction, and (3) eliminating medical errors—and three productivity related goal items—(1) maximizing unit efficiency, (2) delivering cost efficient patient care, and (3) meeting productivity target.

---

[6] We computed this value on the basis of archival revenue data. A .30 standard deviation change in revenue corresponds roughly to $2,762 per patient day. Multiplying this by 20 patients and 365 days results in the value noted.

[7] The effects are computed on the basis of the marginal effect of quality orientation, given high levels (+2 SD) of productivity orientation.

Next, following Cronbach and Gleser (1953), we computed goal incongruence by estimating the D statistic as follows:

$$D_i = \sqrt{\sum_{k=1}^{3} (X_{ik} - Y_k)^2}, \quad \text{(A1)}$$

where X_{ik} represents the rating scores by employee i on the kth measure and Y_k represents the corresponding rating by the unit manager. Higher values of D_i indicate greater incongruence.

After we calculated the individual D_i, we calculated the mean D_i as the indicator of employee–manager goal incongruence separately for productivity and quality goals for each unit:

$$D = \frac{1}{n}\sum_i D_i, \quad \text{(A2)}$$

where n is the number of employees within each unit.

Analyzing the D scores produced the following statistics:

- Quality goal incongruence: $M = .50$, $SE = .06$.
- Productivity goal incongruence: $M = 2.03$, $SE = .09$.
- Comparison of productivity and quality goal incongruence: $t = 13.12$, $d.f. = 46$, $p < .01$.

REFERENCES

Adamy, Janet (2007), "Starbucks Chairman Says Trouble May Be Brewing," *The Wall Street Journal*, (February 24), A4.

Aiken, Linda H., Julie Sochalski, and Eileen T. Lake (1997), "Studying Outcomes of Organizational Change in Health Services," *Medical Care*, 35 (November), NS6–NS18.

Anderson, Eugene W., Claes Fornell, and Roland T. Rust (1997), "Customer Satisfaction, Productivity, and Profitability: Differences Between Goods and Services," *Marketing Science*, 16 (2), 129–45.

Ante, Spencer E. and Ira Sager (2002), "IBM's New Boss," *BusinessWeek*, (February 11), 66–71.

Argote, Linda (1982), "Input Uncertainty and Organizational Coordination in Hospital Emergency Units," *Administrative Science Quarterly*, 27 (September), 420–34.

Bailey, Jeff (2007), "Bumped Fliers and No Plan B," *The New York Times*, (May 30), A1, C4.

Bandura, Albert and Daniel Cervone (1986), "Differential Engagement of Self-Reactive Influences in Cognitive Motivation," *Organizational Behavior and Human Decision Processes*, 38 (August), 92–113.

Baron, James N. and David M. Kreps (1999), "Consistent Human Resource Practices," *California Management Review*, 41 (Spring), 29–53.

Bateson, John E.G. (1985), "Retailing and Services Marketing: Friends or Foes?" *Journal of Retailing*, 61 (Winter), 11–13.

Batt, Rosemary (1999), "Work Organization, Technology, and Performance in Customer Service and Sales," *Industrial & Labor Relations Review*, 52 (July), 539–64.

Bayus, Barry L., Gary Erickson, and Robert Jacobson (2003), "The Financial Rewards of New Product Introductions in the Personal Computer Industry," *Management Science*, 49 (February), 197–210.

Beehr, Terry A. (1976), "Perceived Situational Moderators of the Relationship Between Subjective Role-Ambiguity and Role-Strain," *Journal of Applied Psychology*, 61 (February), 35–40.

Begat, Ingrid, Bodil Ellefsen, and Elisabeth Severinsson (2005), "Nurses' Satisfaction with Their Work Environment and the Outcomes of Clinical Nursing Supervision on Nurses' Experiences of Well-Being: A Norwegian Study," *Journal of Nursing Management*, 13 (May), 221–30.

Boulding, William and Richard Staelin (1995), "Identifying Generalizable Effects of Strategic Actions on Firm Performance: The Case of Demand-Side Returns to R&D," *Marketing Science*, 14 (Summer), 222–36.

Bowen, David E. and Gareth R. Jones (1986), "Transaction Cost Analysis of Service Organization-Customer Exchange," *Academy of Management Review*, 11 (April), 428–41.

——— and Edward E. Lawler III (1992), "The Empowerment of Service Workers: What, Why, How, and When," *Sloan Management Review*, 33 (Spring), 31–39.

Bucklin, Louis P. (1978), "Research in Productivity Measurement for Marketing Decisions," *Research in Marketing*, 1, 1–22.

Campion, Michael A., Gina J. Medsker, and A. Catherine Higgs (1993), "Relations Between Work Group Characteristics and Effectiveness: Implications for Designing Effective Work Groups," *Personnel Psychology*, 46 (December), 823–50.

Carey, Susan and Alex Frangos (2005), "Airlines, Facing Cost Pressure, Outsource Crucial Safety Tasks," *The Wall Street Journal*, (January 21), A1, A5.

Cherrington, David J., Sheri J. Bischoff, W. Gibb Dyer, and Eric G. Stephan (2001), *Organizational Effectiveness*. Hoboken, NJ: John Wiley & Sons.

Cooper, Robert and Michael Foster (1971), "Sociotechnical Systems," *American Psychologist*, 26 (May), 467–74.

Cronbach, Lee J. and Goldine C. Gleser (1953), "Assessing Similarity Between Profiles," *Psychological Bulletin*, 50 (November), 456–73..

Cummings, Thomas G. (1978), "Self-Regulating Work Groups: A Socio-Technical Synthesis," *Academy of Management Review*, 3 (March), 625–34.

Day, George S. and Robin Wensley (1988), "Assessing Advantage: A Framework for Diagnosing Competitive Superiority," *Journal of Marketing*, 52 (April), 1–20.

De Jong, Ad, Ko De Ruyter, and Jos Lemmink (2004), "Antecedents and Consequences of the Service Climate in Boundary-Spanning Self-Managing Service Teams," *Journal of Marketing*, 68 (April), 18–35.

Deci, Edward L. and Richard M. Ryan (1987), "The Support of Autonomy and the Control of Behavior," *Journal of Personality and Social Psychology*, 53 (December), 1024–37.

Durham, Cathy C. and Kathryn M. Bartol (2000), "Pay for Performance," in *Blackwell Handbook of Principles of Organizational Behavior*, Edwin Locke, ed. Malden, MA: Blackwell Publishing, 150–65.

Eisenhardt, Kathleen M. (1985), "Control: Organizational and Economic Approaches," *Management Science*, 31 (February), 134–49.

Fennell, M.L. and J.A. Alexander (1993), "Perspectives on Organizational-Change in the United-States Medical-Care Sector," *Annual Review of Sociology*, 19, 89–112.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Gatignon, Hubert and Jean-Marc Xuereb (1997), "Strategic Orientation of the Firm and New Product Performance," *Journal of Marketing Research*, 34 (February), 77–90.

Gladstein, Deborah L. (1984), "Groups in Context: A Model of Task Group Effectiveness," *Administrative Science Quarterly*, 29 (December), 499–517.

Greene, William H. (2007), *Econometric Analysis*, 6th ed. Englewood Cliffs, NJ: Prentice Hall.

Gregory, Kathleen L. (1983), "Native-View Paradigms: Multiple Cultures and Culture Conflicts in Organizations," *Administrative Science Quarterly*, 28 (September), 359–76.

Grönroos, Christian and Katri Ojasalo (2004), "Service Productivity: Towards a Conceptualization of the Transformation of Inputs into Economic Results in Services," *Journal of Business Research*, 57 (April), 414–23.

Hackman, J. Richard and Greg R. Oldham (1976), "Motivation Through the Design of Work: Test of a Theory," *Organizational Behavior & Human Performance*, 16 (August), 250–79.

Hambrick, Donald C. and Phyllis A. Mason (1984), "Upper Echelons: The Organization as a Reflection of Its Top Managers," *Academy of Management Review*, 9 (April), 193–206.

Harter, James K., Frank L. Schmidt, and Theodore L. Hayes (2002), "Business-Unit-Level Relationship Between Employee Satisfaction, Employee Engagement, and Business Outcomes: A Meta-Analysis," *Journal of Applied Psychology*, 87 (April), 268–79.

Heskett, James L., Earl W. Sasser Jr., and Leonard A. Schlesinger (1997), *The Service Profit Chain*. New York: The Free Press.

Hindo, Brian (2007), "At 3M, a Struggle Between Efficiency and Creativity," *BusinessWeek*, (June 11), 8.

Hollenbeck, John R. and Charles R. Williams (1987), "Goal Importance, Self-Focus, and the Goal-Setting Process," *Journal of Applied Psychology*, 72 (April), 204–211.

House, Robert J. and John R. Rizzo (1972), "Toward the Measurement of Organizational Practices: Scale Development and Validation," *Journal of Applied Psychology*, 56 (October), 388–96.

Jacobson, Robert (1990), "Unobservable Effects and Business Performance," *Marketing Science*, 9 (Winter), 74–85.

Kernan, Mary C. and Robert G. Lord (1990), "Effects of Valence, Expectancies, and Goal-Performance Discrepancies in Single and Multiple Goal Environments," *Journal of Applied Psychology*, 75 (April), 194–203.

Klein, Howard J. (1989), "An Integrated Control Theory Model of Work Motivation," *Academy of Management Review*, 14 (April), 150–72.

Kluger, Avraham N. and Angelo DeNisi (1996), "The Effects of Feedback Interventions on Performance: A Historical Review, A Meta-Analysis, and a Preliminary Feedback Intervention Theory," *Psychological Bulletin*, 119 (March), 254–84.

Landro, Laura (2004), "Report Card to Rank Hospitals on Safety," *The Wall Street Journal*, (April 22), D1, D2.

Langfred, Claus W. and Neta A. Moye (2004), "Effects of Task Autonomy on Performance: An Extended Model Considering Motivational, Informational, and Structural Mechanisms," *Journal of Applied Psychology*, 89 (December), 934–45.

Larkin, T.J. and Sandar Larkin (1996), "Reaching and Changing Frontline Employees," *Harvard Business Review*, 74 (May–June), 95–104.

Levy, Jonathan and Harold Pashler (2001), "Is Dual-Task Slowing Instruction Dependent?" *Journal of Experimental Psychology: Human Perception and Performance*, 27 (August), 862–69.

Locke, Edwin A. and G.P. Latham (1990), *A Theory of Goal Setting and Task Performance*. Englewood Cliffs, NJ: Prentice Hall.

Lytle, Richard S., Peter W. Hom, and Michael P. Mokwa (1998), "SERV*OR: A Managerial Measure of Organizational Service-Orientation," *Journal of Retailing*, 74 (Fall), 455–89.

Mankins, Michael C. and Richard Steele (2005), "Turning Great Strategy into Great Performance," *Harvard Business Review*, 83 (July–August), 65–72.

Mills, Peter K. and Gerardo R. Ungson (2003), "Reassessing the Limits of Structural Empowerment: Organizational Constitution and Trust as Controls," *Academy of Management Review*, 28 (January), 145–53.

Mittal, Vikas, Eugene W. Anderson, Akin Sayrak, and Pandu Tadikamalla (2005), "Dual Emphasis and the Long-Term Financial Impact of Customer Satisfaction," *Marketing Science*, 24 (Fall), 544–55.

Mullen, Brian and Carolyn Copper (1994), "The Relation Between Group Cohesiveness and Performance: An Integration," *Psychological Bulletin*, 115 (March), 210–27.

Noble, Charles and Michael Mokwa (1999), "Implementing Marketing Strategies: Developing and Testing a Managerial Theory," *Journal of Marketing*, 63 (October), 57–74.

Oliva, Rogelio and John D. Sterman (2001), "Cutting Corners and Working Overtime: Quality Erosion in the Service Industry," *Management Science*, 47 (July), 894–914.

Osman, Allen and Cathleen M. Moore (1993), "The Locus of Dual-Task Interference: Psychological Refractory Effects on Movement-Related Brain Potential," *Journal of Experimental Psychology: Human Perception and Performance*, 19 (June), 1292–1312.

Pashler, Harold E. (1998), *The Psychology of Attention*. Cambridge, MA: MIT Press.

Rubinstein, Joshua S., David E. Meyer, and Jeffrey E. Evansy (2001), "Executive Control of Cognitive Processes in Task Switching," *Journal of Experimental Psychology: Human Perception and Performance*, 27 (August), 763–97.

Ruekert, Robert W. (1992), "Developing a Market Orientation: An Organizational Strategy Perspective," *International Journal of Research in Marketing*, 9 (Summer), 225–45.

Rust, Roland T., Christine Moorman, and Peter R. Dickson (2002), "Getting Return on Quality: Revenue Expansion, Cost Reduction, or Both?" *Journal of Marketing*, 66 (October), 7–24.

Schneider, Benjamin (1987), "The People Make the Place," *Personnel Psychology*, 40 (Fall), 437–53.

———, Paul J. Hanges, Brent D. Smith, and Amy Nicole Salvaggio (2003), "Which Comes First: Employee Attitudes or Organizational Financial and Market Performance?" *Journal of Applied Psychology*, 88 (October), 836–50.

Siguaw, Judy A., Gene Brown, and Robert E. Widing (1994), "The Influence of the Market Orientation of the Firm on Sales Force Behavior and Attitudes," *Journal of Marketing Research*, 31 (February), 106–116.

Simons, Robert (2000), *Performance Measurement and Control System for Implementing Strategy*. Englewood Cliffs, NJ: Prentice Hall.

Singh, Jagdip (1998), "Striking a Balance in Boundary-Spanning Positions: An Investigation of Some Unconventional Influences of Role Stressors and Job Characteristics on Job Outcomes of Salespeople," *Journal of Marketing*, 62 (July), 69–86.

——— (2000), "Performance Productivity and Quality of Frontline Employees in Service Organizations," *Journal of Marketing*, 64 (April), 15–34.

Stogdill, Ralph M. (1972), "Group Productivity, Drive, and Cohesiveness," *Organizational Behavior and Human Performance*, 8 (August), 26–43.

Summer, Jane and Judith Townsend-Rocchiccioli (2003), "Why Are Nurses Leaving Nursing?" *Nursing Administration Quarterly*, 27 (April), 164–71.

Taylor, M. Susan, Cynthia D. Fisher, and Daniel R. Ilgen (1984), "Individuals' Reactions to Performance Feedback in Organizations: A Control Theory Perspective," in *Research in Personnel and Human Resources Management*, Vol. 2, Kenneth Rowland and Gerald Ferris, eds. Greenwich, CT: JAI Press, 81–124.

Terborg, James R., Carl Castore, and John A. DeNinno (1976), "A Longitudinal Field Investigation of the Impact of Group Composition on Group Performance and Cohesion," *Journal of Personality and Social Psychology*, 34 (May), 782–90.

——— and Howard E. Miller (1978), "Motivation, Behavior, and Performance: A Closer Examination of Goal Setting and Monetary Incentives," *Journal of Applied Psychology*, 63 (February), 29–39.

Voss, Glenn B. and Zannie G. Voss (2000), "Strategic Orientation and Firm Performance in an Artistic Environment," *Journal of Marketing*, 64 (January), 67–83.

Weinberg, Dana Beth (2003), *Code Green: Money-Driven Hospitals and the Dismantling of Nursing*. Ithaca, NY: ILR Press.

Zeithaml, Valarie A. and Mary Jo Bitner (1996), *Services Marketing*. New York: McGraw-Hill.

Survey Michigan

Courses in Survey Research Techniques

June 2 – July 25, 2008

61st Annual Summer Institute in Survey Research Techniques

Survey Research Center, Institute for Social Research, University of Michigan

For more information: Phone (734) 764-6595, Toll-Free (877) 880-9389 **www.isr.umich.edu/src/si**

Tripat Gill

Convergent Products: What Functionalities Add More Value to the Base?

Convergence in the electronics sector has enabled the addition of disparate new functionalities to existing base products (e.g., adding mobile television to a cell phone or Internet access to a personal digital assistant). This research investigates the role of two factors—(1) the goal congruence between the added functionality and the base and (2) the nature of the base product (utilitarian versus hedonic)—on the evaluation of such convergent products (CPs). The author proposes that the evaluation of CPs with a utilitarian versus hedonic base is subject to an asymmetric additivity effect. Specifically, whereas CPs with a utilitarian base gain more from adding an incongruent, hedonic functionality than a congruent, utilitarian one, CPs with a hedonic base gain less from an incongruent, utilitarian addition than a congruent, hedonic one. This asymmetry is because hedonic additions enhance the pleasure of using a utilitarian base, whereas utilitarian additions may dilute the existing hedonic image of a hedonic base. The moderating role of prior ownership of the base of a CP is also explored. The author proposes that the effects of goal congruence are stronger for owners than for nonowners, but only for CPs with a hedonic base, not for those with a utilitarian base. The author verifies the proposed effects in an experimental study conducted with a large-scale, representative sample of the target market population. Further research on other (moderating) factors affecting the evaluation of CPs is also suggested.

Keywords: convergence, new products, functionalities, consumption goals, utilitarian, hedonic

Convergence is a dominant paradigm in the contemporary high-technology electronics sector—specifically, the intersection of computers, communications, and consumer electronics industries (Yoffie 1997). Such convergence has allowed the introduction of seemingly disparate functionalities into existing products (e.g., the ability to watch television on a cell phone, accessing the Internet on a personal digital assistant [PDA]). Products that are formed by adding a new functionality (from another category) to an existing base product are referred to as "convergent products" (CPs).[1] For example, MP3 music phones add the ability to listen to MP3 music on a cell phone, and Sony's PSP (PlayStation Portable) adds the ability to watch movies and to access the Internet on a handheld gaming device, to name but a few recent examples of CPs. An important managerial issue for such CPs would be determining what kind of new functionalities should be added to a given base product. For example, a manufacturer of PDAs may wonder whether it is a good idea to add mobile television to a PDA, which would mean adding a new functionality that is dissimilar to the existing functions of the PDA. Conversely, it might be better to add a new functionality that is similar to the existing functionalities of a PDA (e.g., adding more information and organizing abilities, such as electronic Yellow Pages). Similarly, a manufacturer of MP3 players (e.g., Apple's iPod) may debate the addition of satellite radio (a functionality congruent with the existing capabilities) versus adding information-associated functions, such as electronic newscasts or weather forecasts (i.e., incongruent new functionalities).

A way to conceptualize the nature of additions in CPs is based on the utilitarian versus hedonic consumption goals associated with the base product and the added functionality. Consumers are known to have both utilitarian and hedonic considerations when evaluating products and functionalities (Batra and Ahtola 1990; Dhar and Wertenbroch 2000). Whereas hedonic goals (and value) are associated with experiential consumption, pleasure, and excitement (e.g., the pleasure experienced through listening to music on an MP3 player), utilitarian goals (and value) are related to more instrumental/practical considerations (e.g., the convenience of using a PDA to keep appointments and contact addresses) (Batra and Ahtola 1990; Hirschman and Holbrook 1982; Voss, Spangenberg, and Grohmann 2003). In the context of CPs, adding new functionalities to an existing base product could result in additions that are goal congruent or incongruent with respect to their utilitarian versus hedonic associations. For example, adding a utilitarian (hedonic) functionality to a utilitarian (hedonic) base would

---

[1] In the current context, new functionality refers to functionalities that are new to the product to which they are added, though they may not be new to the world (e.g., mobile television is not new to the world but is new to the product when added to a PDA).

Tripat Gill is Assistant Professor of Marketing, Faculty of Business and Information Technology, University of Ontario Institute of Technology (e-mail: tripat.gill@uoit.ca). The author gratefully acknowledges the funding received for this research from the Social Sciences and Humanities Research Council of Canada. The author also thanks the three anonymous *JM* reviewers for their guidance and helpful comments on previous versions of this manuscript.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

be goal congruent (e.g., adding electronic Yellow Pages to a PDA, adding satellite radio to an MP3 player), whereas adding a hedonic (utilitarian) functionality to a utilitarian (hedonic) base would be considered goal incongruent (e.g., adding satellite radio to a PDA, adding electronic Yellow Pages to an MP3 player). Table 1 depicts some recent examples of CPs that correspond to these congruent versus incongruent additions to utilitarian versus hedonic base products. Given the recent proliferation of such products, a relevant research issue is whether the nature of additions (goal congruent versus incongruent) affects the evaluation (and success) of CPs. The current research addresses this issue and proposes that goal congruence affects the incremental value of an added functionality in CPs. Whereas congruent functionalities are associated with diminishing incremental value, incongruent functionalities result in an asymmetric effect in CPs with a utilitarian versus hedonic base. Specifically, adding an incongruent, hedonic functionality to a utilitarian base product is valued more than adding an incongruent, utilitarian functionality to a hedonic base one.

The principle of diminishing marginal utility on goals suggests that when a functionality with congruent goals is added to a base product, it will be assimilated, and its incremental value will be subject to diminishing returns (see Nowlis and Simonson 1996). For example, adding electronic Yellow Pages to a PDA (or satellite radio to an MP3 player), both of which are congruent additions, adds utilitarian (hedonic) value to an already utilitarian (hedonic) base and thus will be subject to diminishing marginal utility. Conversely, when an incongruent functionality is added in a CP, there are two possibilities of contrast: (1) A new functionality associated with hedonic goals is added to a utilitarian base product (e.g., adding satellite radio to a PDA), or (2) a functionality associated with utilitarian goals is added to a hedonic base product (e.g., adding electronic Yellow Pages to an MP3 player). It is proposed that the former addition will be positively contrasted, whereas the latter will be negatively contrasted. This is because, with reference to a utilitarian base, an added hedonic functionality will enhance the value of the base, and the CP will be perceived as more fun/pleasurable to use than the original base product. Conversely, an added utilitarian functionality could dilute the value of a hedonic base, and the CP could be perceived as less fun/pleasurable to use than the original base product. The former enhancing effect on a utilitarian base is because hedonic functionalities are more pleasurable (Hirschman and Holbrook 1982), are easier to imagine (MacInnis and Price 1987), and are associated with a spontaneous positive affect (Shiv and Fedorikhin 1999), whereas the diluting effect on a hedonic base is because utilitarian functionalities are associated with practical/instrumental goals, which are perceived as less fun/pleasurable than hedonic consumption goals (Hirschman and Holbrook 1982); this negative correlation is akin to the lay belief that "utilitarian is less fun." When incongruent functionalities are added to a utilitarian versus hedonic base, this is the "asymmetric additivity effect." In summary, for CPs with a utilitarian base, adding an incongruent, hedonic functionality offers more incremental value than adding a congruent, utilitarian functionality, whereas for CPs with a hedonic base, adding an incongruent, utilitarian functionality offers less incremental value than adding a congruent, hedonic functionality.

These proposed effects are tested and confirmed in an experimental study with a representative sample of the target market population. In addition, the study also explores the moderating role of prior ownership of the base on the evaluation of CPs (i.e., whether owners versus nonowners of the base product evaluate CPs differently). Because owners have a higher involvement in and liking for the base product than nonowners, they will be more sensitive to changes in the perceived image of the base (Kirmani, Sood, and Bridges 1999)—and especially to the dilution effect—when incongruent, utilitarian functionalities are added to a hedonic base (e.g., when electronic Yellow Pages are added to an MP3 player). Accordingly, the effects of congruent versus incongruent additions will be stronger for owners than for nonowners and more so for CPs with a hedonic base than for those with a utilitarian one. The study also confirms the moderating effects of prior ownership. The results obtained from this empirical study have implications for managers for both the development of new CPs and the targeting of existing versus new customers (i.e., owners versus nonowners of the base). This is particularly relevant in the current context of convergence in the electronics and communications industry, which has led to a significant increase in CPs.

The rest of the article is organized as follows: In the subsequent section, the existing literature relevant to the addition of new functionalities in products is reviewed. Then, a theoretical model is developed, and hypotheses are proposed for the evaluation of CPs with a utilitarian versus hedonic base. These hypotheses are tested in an experimental study among a sample of the general population. Finally, the theoretical and managerial implications of the findings are discussed, followed by the limitations and avenues for further research.

TABLE 1
Examples of CPs with New Functionalities Added to a Utilitarian or Hedonic Base

Convergent Product	Base Product	Added Functionality	Nature of Addition (Functionality + Base)
Garmin's iQue 3000	PDA	Global positioning system	Congruent (utilitarian + utilitarian)
Blackberry Pearl	Handheld e-mail device	Camera and MP3 music	Incongruent (hedonic + utilitarian)
Video iPod	Portable MP3 music player	Video capability	Congruent (hedonic + hedonic)
Web television	Television	Internet access	Incongruent (utilitarian + hedonic)

Literature Review and Theoretical Model

This research investigates the incremental value of adding new functionalities in CPs. Specifically, the effect of goal congruence is explored in CPs that add utilitarian versus hedonic functionalities to existing utilitarian versus hedonic base products. The theory of assimilation/contrast effects and the principles behind adding new features/functionalities are relevant to the current research. The existing literature on these two mechanisms is first reviewed briefly, and a theoretical model is proposed with specific research hypotheses.

Assimilation/Contrast Effects When New Functionalities Are Added to Products

Assimilation versus contrast is a robust psychological phenomenon observed in the judgment of new stimuli in the assessment of both people and objects (Herr, Sherman, and Fazio 1983; Schwarz and Bless 1992). Specifically, a new stimulus can be either assimilated in the direction of a context or contrasted away from it, depending on factors such as the degree of feature overlap between the context and the stimulus to be evaluated (Herr, Sherman, and Fazio 1983). In the context of product perceptions, pioneering new products are judged more positively than "follower" products because the former are contrasted with other products (and thus perceived as differentiated), whereas the latter are assimilated and perceived as similar to the pioneer (Carpenter and Nakamoto 1989). In a similar vein, Gatignon and Xuereb (1997) find that the similarity of an innovation to its competitors is negatively correlated with financial performance and results in lower competitive advantage to the innovator. In the context of new products, Meyers-Levy and Tybout (1989) find that adding moderately incongruous features is favored over congruous ones (in which "feature incongruity" is defined in terms of the perceived distance between the feature and the product in a hierarchical category structure). Finally, and more pertinent to the current context, Ziamou and Ratneshwar (2003) find assimilation/contrast effects when a new functionality is added to a product and is compared with another existing functionality. For example, they find that when Anytime Entertainment (a hypothetical new functionality that allows receipt of entertainment programs—movies, music, and so forth—through cable) is compared with cable television, participants evaluate it more favorably when it is introduced in a product that is atypical of cable television (e.g., in a personal computer [PC], which leads to contrast) than when it is introduced in a product typical of it (e.g., in a television, which leads to assimilation). Whereas the previous work has investigated the assimilation/contrast effects when new features are added to a product, the effects have been proposed primarily on the basis of feature similarity (typicality), but the role of goal congruence has not been explored.[2] More important, the assimilation/contrast effects resulting from the addition of new functionalities to an existing utilitarian or hedonic base have not been examined.

Functionalities (and products) enable consumers to engage in goal-directed actions (Ratneshwar et al. 1999; Ziamou and Ratneshwar 2003), and consumption goals can be broadly defined as utilitarian or hedonic, depending on the value derived from achieving these goals. Whereas hedonic value is associated with achieving experiential goals (e.g., the pleasure of listening to music on an MP3 player), utilitarian value is derived from achieving instrumental/functional goals (e.g., the convenience of using a PDA to keep appointments and contact addresses) (Batra and Ahtola 1990; Hirschman and Holbrook 1982). With this typology, products (and functionalities) can be classified as hedonic versus utilitarian (Batra and Ahtola 1990; Voss, Spangenberg, and Grohmann 2003). Thus, an MP3 player would be associated with hedonic goals, and a PDA would be associated with utilitarian goals, because consumers who use these products derive relatively more hedonic value (than utilitarian) from the former and more utilitarian value (than hedonic) from the latter. Similarly, a functionality such as electronic Yellow Pages would be associated with utilitarian goals, whereas satellite radio would be associated with hedonic goals.

In the context of CPs, the goal congruence between the added functionality and the base product is defined as the extent to which the added functionality and the base product are associated with similar/different goals in terms of their utilitarian versus hedonic value.[3] Thus, a PDA with the added functionality of electronic Yellow Pages (both associated with utilitarian goals) and an MP3 player with added satellite radio (both associated with hedonic goals) would be considered goal congruent, whereas a PDA with satellite radio and an MP3 player with electronic Yellow Pages would be considered goal incongruent. A key issue explored here is how consumers perceive the incremental value of CPs that have congruent versus incongruent functionalities added to a base product.

---

[2] Note that Ziamou and Ratneshwar (2003) define "typicality" of the added functionality in terms of the goal-derived category with which it is associated (e.g., cable programming is typical of televisions). However, their main focus is to investigate the impact of comparing these added functionalities with an existing functionality that is typical versus atypical of the product to which it is added. They find that when explicit comparisons are made, the new functionalities are favored when added to atypical products rather than to typical ones. The authors only explore the effect of explicit comparisons but not the baseline effect of adding new functionalities to products. Furthermore, they do not examine the effect of utilitarian versus hedonic goals associated with the added functionality and the base product.

[3] Goals are defined as abstract benefits that consumers seek in specific consumption situations (Huffman and Houston 1993), and consumers can achieve utilitarian or hedonic value by achieving these goals. In the current context, goal congruence implies that consumers seek similar (or dissimilar) benefits and value from the added functionality and the base product. Although a given functionality and a base may be associated with similar benefits (goals) but may lead to different (utilitarian or hedonic) value, in the current context, it is assumed that both the benefits (goals) and the associated value are congruent (or incongruent). This similarity in benefits and value was incorporated in the empirical measure of goal congruence used in the experimental study reported subsequently.

Evaluation of CPs with Congruent/Incongruent Functionalities Added to a Utilitarian Versus Hedonic Base

It is proposed that assimilation versus contrast effects will be initiated contingent on the goal congruence between the added functionality and the base product of a CP. Specifically, when a functionality with congruent goals is added to a base (i.e., both associated with utilitarian or hedonic goals), the overlap in the underlying consumption goals will enable the new functionality to be assimilated into the base. Conversely, when a functionality with incongruent goals is added to a base (i.e., one associated with utilitarian goals and the other with hedonic goals), the lack of overlapping goals will lead the CP to be contrasted with the existing base product. The ensuing assimilation versus contrast processes will affect how consumers add the value (or utility) associated with the functionality and the base product of the CP. These differential additive effects will manifest in two ways.

First, there is a diminishing marginal utility of congruent functionalities. According to the principle of diminishing marginal utility, the value curves for goods are concave, and the incremental value of adding features reduces with the overall value of the product (Nowlis and Simonson 1996). This principle is derived from Weber's law that the sensitivity to increases along a dimension reduces as the magnitude of the dimension increases (in the context of sensory perception in psychophysics, see Stevens 1986). In the current context, when a congruent, utilitarian (hedonic) functionality is added to a utilitarian (hedonic) base product, it is assimilated, and the CP will be subject to diminishing returns because of similar utilitarian (hedonic) goals. For example, when electronic Yellow Pages is added to a PDA (thus sharing similar utilitarian goals related to contact information, such as addresses, phone numbers, and so forth) or when satellite radio is added to an MP3 player (thus sharing similar hedonic goals related to listening to music), the evaluation of the CP will be subject to diminishing returns (i.e., the incremental value of the added functionality decreases as the prior value of the base increases). Thus, the overall value of such CPs (with congruent functionalities added to a base) is subadditive (i.e., the whole is less than the sum of its parts).

Second, there is an asymmetric additivity effect of incongruent functionalities. Incongruent functionalities are contrasted with the base product, and the resultant additivity of the two components depends on the valence of the contrast (i.e., whether the functionality is positively or negatively contrasted with respect to the base). Previous research has shown that people evaluate options with respect to their reference point, and they display loss aversion in their evaluations (i.e., a loss is weighted more than an equivalent gain from the reference point) (Tversky and Kahneman 1991). Furthermore, the loss aversion is larger for hedonic attributes than equally valued utilitarian ones (Dhar and Wertenbroch 2000). For example, when market prices are controlled for, owners of relatively hedonic cars are more wary of selling their cars (i.e., losing the hedonic benefits) than owners of relatively utilitarian cars, and accordingly, the former demands a higher price to forfeit their cars than the latter (see the field survey in Dhar and Wertenbroch 2000).

For CPs with a utilitarian base, adding an incongruent, hedonic functionality will be positively contrasted and perceived as a gain of hedonic value with respect to the base. This is because hedonic functionalities are perceived as more pleasurable/exciting than utilitarian ones (Hirschman and Holbrook 1982), and thus adding them means providing more pleasure/fun in using the utilitarian base product. Furthermore, because hedonic functionalities are more sensorial and image evoking (MacInnis and Price 1987), they can have a stronger influence on evaluations than the existing utilitarian attributes of the base (Keller and McGill 1994), and because they are associated with pleasure/fun, they can elicit spontaneous (positive) affective reactions (Shiv and Fedorikhin 1999).[4] Thus, when a hedonic functionality is added to a utilitarian base product (e.g., adding satellite radio to a PDA), it will be perceived as enhancing the (hedonic) value of the utilitarian base. Accordingly, the overall value of such CPs is additive (i.e., the whole should equal the sum of its parts).[5]

Conversely, for CPs with a hedonic base, adding an incongruent, utilitarian functionality (e.g., adding electronic Yellow Pages to an MP3 player) will be negatively contrasted and perceived as a loss of hedonic value with reference to the base. This is because utilitarian functionalities are associated with more practical/instrumental goals that are perceived as less pleasurable/exciting than hedonic goals (corresponding to the lay belief that utilitarian is less fun) (Hirschman and Holbrook 1982).[6] Furthermore, although the added utilitarian functionality can result in a gain in utilitarian value, the associated loss in hedonic value of the base product will be weighted more heavily (Dhar and Wertenbroch 2000; Tversky and Kahneman 1991). Thus, the overall value of such a CP (an incongruent, utilitarian functionality added to a hedonic base) will be unadditive (or subtractive).[7]

---

[4] Although Shiv and Fedorikhin (1999) show that the affective reactions have a stronger influence (than cognitions) only when processing resources are limited, Keller and McGill (1994) argue that easily imaginable attributes carry more weight under conditions of high elaboration. In either case, because of spontaneous evaluations or with more elaboration, the added hedonic functionality should have a greater influence on the evaluations of such CPs than the existing utilitarian functions of the base product.

[5] Here, it is assumed that the added hedonic functionality will not interfere with the existing utilitarian functions of the base product of the CP.

[6] The proposed negative correlation between utilitarian functionalities and their hedonic value is based on the lay belief that utilitarian is less fun. Conversely, the opposite belief for hedonic functionalities (i.e., "hedonic is less useful") may not be as widely held. This asymmetry may be due to the increased emphasis on experiential consumption (i.e., hedonic value) in highly developed economies, such as those of the United States, Canada, and Europe (see Pine and Gilmore 1998).

[7] Recall that the added utilitarian functionality is assumed to be incongruent with respect to both the goals and the value offered by the base hedonic product. However, there could be cases in which an added utilitarian functionality is related to the goals of the hedonic base product (e.g., adding "wireless music downloading" ability to an MP3 player). Such additions are discussed subsequently in the "Limitations and Further Research" section.

This asymmetry in the evaluation of CPs with an incongruent functionality added to a utilitarian versus hedonic base is called the asymmetric additivity effect.

> H_1: The incremental value of a CP with a hedonic functionality added to a utilitarian base is greater than that with an equally valued utilitarian functionality added to a hedonic base.

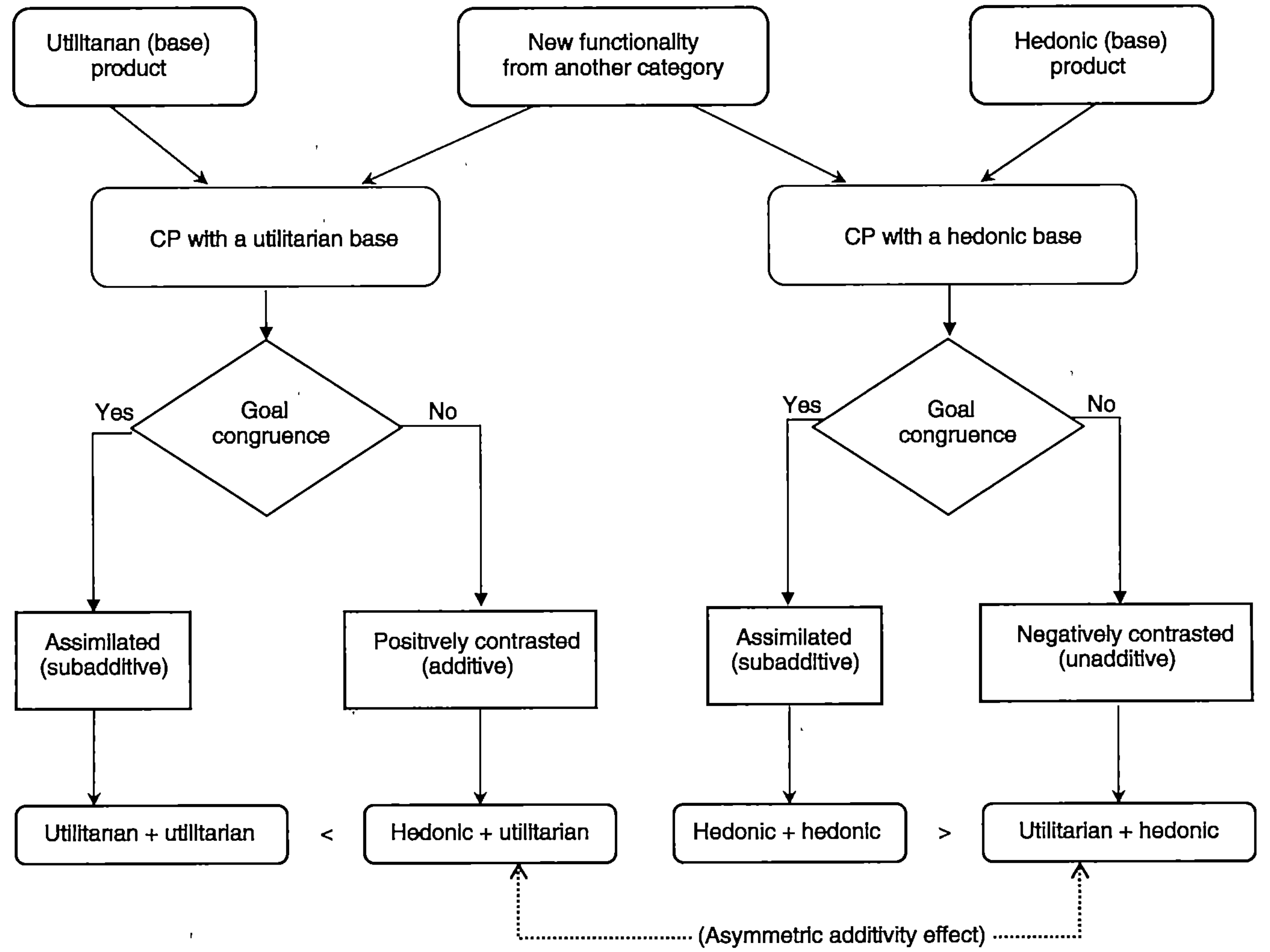
Figure 1 depicts the proposed additive processes in the evaluation of CPs. As is evident from the figure, congruent functionalities are assimilated for both utilitarian and hedonic base products, whereas there is an asymmetry in the evaluation of CPs with incongruent functionalities added to a utilitarian versus hedonic base product. Based on this theoretical model, specific hypotheses can be proposed for CPs with a utilitarian versus hedonic base.

CPs with a utilitarian base. For CPs with a utilitarian base, adding a congruent, utilitarian functionality is subject to diminishing utility on goals, whereas adding an incongruent, hedonic functionality is positively contrasted and is perceived to enhance the base product. Thus, adding a functionality associated with hedonic goals will have more incremental value than adding one associated with utilitarian goals (e.g., a PDA with satellite radio will have more incremental value than a PDA with electronic Yellow Pages). As discussed previously, this is because hedonic functionalities are perceived as more pleasurable/fun (and are associated with positive affect) than utilitarian ones (Hirschman and Holbrook 1982), and thus adding them entails a gain in hedonic value for the utilitarian base product.

> H_2: For CPs with a utilitarian base, adding a new functionality associated with incongruent, hedonic goals has (a) greater incremental value and (b) greater incremental pleasure (hedonic value) than adding an equally valued functionality associated with congruent, utilitarian goals.

CPs with a hedonic base. For CPs with a hedonic base, adding an incongruent, utilitarian functionality will be negatively contrasted with the hedonic base and could dilute its hedonic image, whereas adding a congruent, hedonic functionality will maintain the hedonic image of the base and add positive value, albeit with diminishing returns. Thus, for example, an MP3 player with electronic Yellow

FIGURE 1
Theoretical Model for the Evaluation of CPs with a Utilitarian Versus Hedonic Base


Utilitarian (base) product
New functionality from another category
Hedonic (base) product
CP with a utilitarian base
CP with a hedonic base
Yes
Goal congruence
No
Yes
Goal congruence
No
Assimilated (subadditive)
Positively contrasted (additive)
Assimilated (subadditive)
Negatively contrasted (unadditive)
Utilitarian + utilitarian
<
Hedonic + utilitarian
Hedonic + hedonic
>
Utilitarian + hedonic
(Asymmetric additivity effect)

Pages will have less incremental value than an MP3 player with satellite radio. As discussed previously, this is because utilitarian functionalities are associated with practical/instrumental goals that are perceived as less pleasurable/fun than hedonic goals, and thus adding them implies a loss in hedonic value for a hedonic base product.

H_3: For CPs with a hedonic base, adding a new functionality associated with incongruent, utilitarian goals has (a) less incremental value and (b) less incremental pleasure (hedonic value) than adding an equally valued functionality associated with congruent, hedonic goals.

Prior Ownership Effects in the Evaluation of CPs

There could be potential moderators to the additive processes in CPs, and prior ownership (of the base) is proposed as one such factor. Prior ownership is an important factor of both theoretical and practical interest. Theoretically, ownership moderates the prior value of the base product (i.e., owners value the base more than nonowners, according to the mere ownership effect; see Beggan 1992), and thus owners may be more sensitive to the changes in the image of the base of a CP than nonowners (Kirmani, Sood, and Bridges 1999). Prior ownership also has practical implications; first, it is a commonly used variable for segmentation, and second, it distinguishes the purchase of a CP into an upgrade (for owners of the base) from a new purchase (for nonowners of the base). Because the decision to upgrade is qualitatively different from that of making a new purchase (Okada 2006), prior ownership (of the base) could affect the evaluation of CPs.

Prior ownership can influence the incremental value of a CP through two mechanisms: (1) the perceived sunk cost associated with the base for owners compared with nonowners and (2) the perceived change in image of the base product when a new functionality is added to it. Okada (2006) argues that upgrades (i.e., existing products with enhanced features) are viewed differently by owners and nonowners because the former also consider the costs they have invested in the existing product they own (though it is a sunk cost). However, Okada finds that sunk costs play a role only when the upgrade has similar enhancements, but not when it has dissimilar ones (because it is easier to write off the sunk costs when the upgrade is dissimilar to the existing product). The similarity of an enhancement has been defined in terms of its alignability (Markman and Gentner 1993); namely, enhancements are similar when they are made to existing (alignable) features (e.g., increasing the storage capacity of an MP3 player from 4 GB to 8 GB) and are dissimilar when they are made by adding new (i.e., nonalignable) features (e.g., adding video capability to an MP3 player). By this account, the current research on CPs focuses solely on nonalignable enhancements because only the effect of adding new functionalities is investigated.[8] Because sunk costs are a more important consideration for alignable enhancements than for nonalignable ones (Okada 2006), perceived sunk costs should play a limited role in the prior owners' evaluation of CPs. In contrast, the perceived change in image of the base should be a more important consideration for prior owners.

Ownership is linked to familiarity, knowledge, and involvement in a product category (Kirmani, Sood, and Bridges 1999). In addition, ownership enhances the liking for the owned product (Beggan 1992) and can lead to owners developing a symbolic relationship to their possessions (Belk 1988). Such an attachment can affect how owners evaluate the extensions of a product/brand. For example, Kirmani, Sood, and Bridges (1999) find that when a prestigious car brand (e.g., BMW) is extended downward to introduce a lower-priced model, prior owners are more unfavorable to the extension than nonowners. The theoretical explanation is that because owners are more involved with the brand and, thus, more likely to maintain the core image of the brand, they will react unfavorably to a downward extension of a prestige brand, whereas nonowners will react favorably (Kirmani, Sood, and Bridges 1999). However, this ownership effect is not found for functional brands (e.g., if Honda were to introduce a lower-priced car). In a similar vein, in the context of CPs (which are product extensions/enhancements), prior owners of the base should be more concerned about the perceived change in image of the base than nonowners.[9] Such concern should also be more likely for owners of a hedonic base than for owners of a utilitarian one.

As proposed previously, the perceived image of the base product of a CP can change when new functionalities are added to it. For example, it is proposed that adding a hedonic functionality to a utilitarian base will enhance the image of the base (making it more fun to use), whereas adding a (incongruent) utilitarian functionality to a hedonic base will dilute the base image. Given that owners have a stronger motivation to maintain the image of the base product than nonowners (Kirmani, Sood, and Bridges 1999), they will be more sensitive to the dilution in the image of a (hedonic) base than nonowners. However, no such difference is expected in the case of a utilitarian base, which is enhanced by adding an incongruent functionality. This is because (1) to the extent that hedonic products are more symbolic and experiential (Hirschman and Holbrook 1982; MacInnis and Price 1987), owners are more likely to form an attachment/relationship to hedonic possessions than to utilitarian ones, and (2) loss aversion is higher for hedonic attributes (Dhar and Wertenbroch 2000), the perceived change (dilution) in image of a hedonic base should be weighted more by owners than that of a change in a utilitarian base. Accordingly, for CPs with a hedonic base, adding congruent versus incongruent functionalities will have a greater effect on perceived incremental value for owners of the base than for nonowners. For CPs with a utilitarian base, however, no such difference is expected. Thus:

---

[8] Although adding a congruent functionality to a base product (e.g., adding satellite radio to an MP3 player) can be construed as a similar enhancement, because both share the same consumption goals, it is not an alignable enhancement. This implies augmenting an existing feature (e.g., increasing the memory of an MP3 player). Thus, it is assumed that adding a congruent functionality is a nonalignable enhancement in the base.

[9] Here, "perceived image" refers to the type of product concept associated with the base—namely, a functional concept for a utilitarian base and an experiential concept for a hedonic base.

> H_{4a}: For CPs with a utilitarian base, the effect of goal congruence on the incremental value of the added functionalities is equal for both prior owners and nonowners of the base product.
>
> H_{4b}: For CPs with a hedonic base, the effect of goal congruence on the incremental value of the added functionalities is greater for prior owners of the base product than for nonowners.

These hypothesized ownership effects are proposed on the basis of the assumption that the owners will be more sensitive to any change (dilution) in the image of the hedonic base than nonowners. However, this will not be the case for owners of a utilitarian base (1) because maintaining (and not losing) the image is less of a concern for a utilitarian base than for a hedonic one and (2) because the utilitarian image is not likely to be diluted by adding an incongruent, hedonic functionality. Thus:

> H_{5a}: For CPs with a utilitarian base, the effect of added functionalities on the perceived change in image of the base is equal for owners and nonowners of the base product.
>
> H_{5b}: For CPs with a hedonic base, the effect of added functionalities on the perceived change in image of the base is greater for owners of the base product than for nonowners.

All the proposed hypotheses pertaining to both the overall effects of goal congruence on CP evaluations (H_1–H_3) and the moderating effects of prior ownership (H_4–H_5) were tested in a survey-based experimental study with a representative sample of the target population. Before the main study, an extensive pretest was conducted to select functionalities that were associated with utilitarian or hedonic goals and were accordingly congruent or incongruent with respect to the chosen base products.

Pretest

The base product categories chosen for the main study were a PDA and an MP3 player because the former is associated with relatively utilitarian goals (i.e., organizing ability, contact addresses, and other information) and the latter is associated with relatively hedonic goals (i.e., listening to music/entertainment). Moreover, these product categories were relatively pure insofar as they were linked to singular functionalities, as opposed to other devices that have multiple functionalities (e.g., cell phones, PCs). The sample for the main study was provided by a market research firm specializing in online research. The firm has a database of more than 100,000 consumers, representative of Internet users in the United States, who have agreed to participate in online surveys in exchange for cash incentives. Note that though this sample is not representative of the general U.S. population, it is appropriate for the current study because it represents the target market for the two base products.

After the selection of the base products and the sample, the objective of the pretest was to select two functionalities such that (1) one functionality was associated with utilitarian goals and the other was associated with hedonic goals; (2) one functionality was (goal) congruent with a PDA and incongruent with an MP3 player, and vice versa for the other; and (3) both functionalities were of equal perceived value. The pretest was conducted with a smaller sample (n = 200, ages 18–54) from the same online panel of respondents that was used for the main study. In addition, the participants were screened such that only those who were aware of the base products could participate.[10]

Measures

The qualified participants responded to three sets of measures pertaining to the two base products (the PDA and the MP3 player) and to the eight functionalities that were chosen as potential additions to these products. The eight functionalities were satellite radio, electronic daily newspaper, electronic dictionary, video games, global positioning system (GPS), electronic Yellow Pages, electronic weather forecasts, and local bus/train schedule.[11] The three sets of measures obtained for the two base products and the eight functionalities were (1) hedonic and utilitarian value, which were measured with two items each, selected from the ten-item scale that Voss, Spangenberg, and Grohmann (2003) developed—namely, "unhelpful/helpful" and "impractical/practical" for utilitarian value (averaged to one measure; r = .79) and "not fun/fun" and "dull/exciting" for hedonic value (averaged to one measure; r = .82) (all scales ranged from 1 to 7); (2) goal congruence, which was measured with two items adopted from Martin and Stewart (2001, Table 1)—namely, "How similar is the goal associated with (base product) and (added functionality)?" and "How similar is the reason for using (base product) and (added functionality)?" (both scales ranged from 1 ["not at all similar"] to 7 ["very similar"] and were averaged to one measure; r = .84)[12]; and (3) overall perceived value, which was measured with the item, "Overall how valuable is the base product (functionality) to you?" (scale ranged from 1 ["not at all"] to 9 ["extremely valuable"]).

---

[10] The two screening questions used were as follows: (1) "Which of the following products are you aware of and know what they are used for: (a) cellular phone, (b) satellite radio, (c) PDA, (d) MP3 music player, (e) global positioning system (GPS) for directions, and (f) none of the above"; and (2) "Which of the following products are you Not aware of and do not know what they are used for: (a) cellular phone, (b) satellite radio, (c) PDA, (d) MP3 music player, (e) GPS for directions, and (f) none of the above." Only respondents who selected the PDA and MP3 music player in Question 1 and did not select the PDA or MP3 music player in Question 2 were allowed to participate in the surveys.

[11] In addition to the screening questions pertaining to the base products, participants were also screened such that only those who were aware of satellite radio and GPS were allowed to participate because these were two functionalities of which the participants may not be well aware. For the other six added functionalities, it was assumed that all participants were equally aware of these because they are fairly commonplace.

[12] These items for goal congruence measured the similarity in benefits (i.e., reasons to use) associated with the base and the added functionalities. As discussed previously (n. 3), this measure was to ensure that the added functionality and the base product were congruent (or incongruent) with respect to both the goals (i.e., benefits or reasons to use) and the value (i.e., utilitarian or hedonic) associated with them.

Results

For the two base products, the PDA was associated with relatively more utilitarian than hedonic value (5.76 versus 5.29; $t_{(200)} = 5.60$, $p = .000$), whereas the opposite was true for the MP3 player (5.06 versus 6.04; $t_{(200)} = -11.84$, $p = .000$). Among the eight functionalities tested, only two met the requirements for all the three criteria for selection, namely, electronic Yellow Pages and satellite radio. Specifically, the three selection criteria were: (1) utilitarian versus hedonic value—the Yellow Pages was associated with relatively more utilitarian than hedonic value (5.52 versus 4.47; $t_{(99)} = 7.23$, $p = .000$), whereas the opposite was true for satellite radio (5.23 versus 5.77; $t_{(99)} = -5.25$, $p = .000$); (2) goal congruence—for the PDA, the Yellow Pages was congruent, and satellite radio was incongruent (4.76 versus 2.82; $t_{(198)} = 8.53$, $p = .000$), whereas for the MP3 player, the Yellow Pages was incongruent, and satellite radio was congruent (2.25 versus 5.42; $t_{(198)} = -15.41$, $p = .000$); and (3) overall perceived value—the Yellow Pages and satellite radio did not differ in their overall perceived value (6.36 versus 6.16; $t_{(198)} = .61$, $p = .541$).

On the basis of this pretest, the Yellow Pages and satellite radio were selected as the functionalities to be added to the two base products (the PDA and the MP3 player); four CPs were created corresponding to the four conditions described in the theoretical model developed previously (see Figure 1). These CPs were (1) the Yellow Pages added to a PDA (YEL-PDA)—that is, a utilitarian functionality added to a utilitarian base (goal congruent); (2) satellite radio added to a PDA (SAT-PDA)—that is, a hedonic functionality added to a utilitarian base (goal incongruent); (3) satellite radio added to an MP3 player (SAT-MP3)—that is, a hedonic functionality and a hedonic base (goal congruent); and (4) the Yellow Pages added to an MP3 Player (YEL-MP3)—that is, a utilitarian functionality added to a hedonic base (goal incongruent).

Main Study

Method

Study design and sample. The study was a 2 (utilitarian or hedonic base product) × 2 (congruent or incongruent added functionality) between-subjects design, and the stimuli used to manipulate these two factors were the four CPs. The study was conducted with a representative sample of the potential target market for the two base products. The sample was provided by a market research firm that has a panel of respondents who participate in online surveys for cash incentives, as described in the pretest. The study was conducted as four separate surveys (one each for each of the four CPs: YEL-PDA, SAT-PDA, SAT-MP3, and YEL-MP3), each of which were administered online through the market research firm's Web site. A total of 20,000 (5000 for each survey) invitations were sent by e-mail to a randomly selected pool of panelists from the database. Each panel member was provided with a unique ID and password, and members were not allowed to complete the survey more than once. This controlled for the possibilities of multiple responses and the occurrence of random "walk-ins" (i.e., participants who randomly end up on the Web site without being invited). In addition, there were two screening questions (same as those used in the pretest), which ensured that only people aware of the base products and the added functionalities (i.e., PDAs, MP3 players, satellite radio, and the Yellow Pages) could participate in the surveys.[13]

The survey was administered until 500 completed responses were obtained for each of the four surveys. Of the 20,000 invitations, 2999 people logged on to participate in the survey, for an overall response rate of 15.0% (15.2% for the survey pertaining to YEL-PDA, 14.8% for SAT-PDA, 14.5% for SAT-MP3, and 15.4% for YEL-MP3). Of the 2999 participants, 999 were screened out (on the basis of their responses to the screening questions), and 1 survey was discarded because of incomplete responses, resulting in 1999 completed surveys (499 for YEL-PDA, and 500 each for SAT-PDA, SAT-MP3, and YEL-MP3). The completed surveys were collected in a span of approximately one week during which the survey was posted on the market research firm's Web site; it was removed after the desired sample was obtained. According to the firm's policy, people who completed the survey were randomly awarded $1 (80%), $2 (10%), or $3 (10%) for their participation. In addition, their names were added to a lottery for $5,000 to be awarded after they completed a specified number of projects (including this survey).

Appendix A provides a demographic distribution of the sample that completed the four surveys, including variables of gender, age, income, and education (for more than 96% of the respondents, the native language was English in each of the surveys). Appendix A also provides a comparative distribution of the U.S. population, according to the data obtained from the market research firm that administered the surveys. As is evident from the comparison, the sample used in the online survey was relatively younger and of a higher education level than the general population. This discrepancy is not an issue for the current study, because the demographic composition is consistent with the likely target market for the two base products.

Procedure. The participants were randomly assigned to one of the four surveys on the basis of the 2 (utilitarian versus hedonic base) × 2 (goal congruence of the added functionality) study design. Each respondent first consented to participate and then qualified on the basis of the two screening questions. They then saw the description of one of the four CPs created by adding a new functionality (Yellow Pages or satellite radio) to an existing base product (the PDA or the MP3 player) (for the description used for two of the four CPs employed in the study—namely, for YEL-PDA and SAT-MP3—see Appendix B). Participants were informed that the CP would soon be launched in the market by a leading manufacturer and that consumers' reactions to the product were being sought. Participants were encouraged to read the product description carefully before moving ahead in the survey. They then responded to several questions pertaining to the CP, the base product, and the added functionality. In addition, the prior ownership of the base product corresponding to the particular CP was measured by asking participants to indicate whether they owned a PDA (an MP3 player). The question specified that the PDA (MP3 player) should be an independent portable device associated only with the functions of a PDA (an MP3 player), not a combination device that had a PDA (an MP3 player) integrated into it. Finally, the participants' demographic information was collected.

---

[13] As noted in the pretest, the screening questions checked for the awareness of the PDA, the MP3 player, and satellite radio, whereas the awareness of the Yellow Pages was assumed. A post hoc measure of prior knowledge of the two added functionalities (on three items, adopted from Mitchell and Dacin [1996], measuring "familiarity with features," "familiarity with usage," and "overall knowledge" [all nine-point scales; $\alpha = .97$]) revealed that the prior knowledge of the Yellow Pages was higher than that for satellite radio (7.29 versus 5.04; $t(1997) = 24.71$, $p = .000$). However, this difference would not affect the predictions made here for the incremental value of the CPs, which are proposed to be based on the goal congruence between the added functionalities and the base product. On the contrary, the higher knowledge for the Yellow Pages (than for satellite radio) should act in its favor in terms of the incremental value added; for example, this would make it a more difficult test for one of the predictions (namely, that SAT-PDA should have a higher incremental value than YEL-PDA).

Dependent measures. The main dependent measure was the perceived incremental value of the CP, compared with the base product to which the new functionality was added. This was measured in two ways: (1) Overall incremental value (OIV) was assessed with the question, "Overall, how much more (or less) valuable was the new product (i.e., the CP) than the base product?" (nine-point scale ranging from –4 ["much less valuable"] to +4 ["much more valuable"], and (2) incremental willingness to pay (IWTP) was assessed with the question, "How much more are you willing to pay for this new product (CP) as compared to the base product?"[14] In addition, the incremental pleasure (i.e., hedonic value) of using the CP was assessed with the question, "How much more pleasure would you feel using this product as compared to the base?" (nine-point scale ranging from –4 ["much less pleasure"] to +4 ["much more pleasure"]). This question was assessed to verify the enhancing versus diluting effect of adding incongruent functionalities to a utilitarian versus hedonic base product, respectively (as proposed in the asymmetric additivity effect).

Other measures. The perceived utilitarian and hedonic value was measured (with the same two-item scales employed in the pretest) for both the base products (r = .76 and .83 for the scales for utilitarian and hedonic value, respectively) and the added functionalities (r = .83 and .86, respectively). Similarly, the goal congruence between the base product and the added functionality was measured on the same two-item scales used previously (r = .81), and the overall perceived value of the base products and the functionalities was measured on the same scale used in the pretest. These measures were used in the manipulation checks and in other analyses. In addition, to verify the prior ownership effects, the perceived change in image of the base products was measured as the change in the existing hedonic (utilitarian) image of the hedonic (utilitarian) base product. The change in hedonic image was measured with two items ("How much more fun [exciting] was the new product [i.e., the CP] than the base?") on seven-point scales ranging from –3 to +3 (r = .88), and the change in utilitarian image was measured with two items ("How much more helpful [practical] was the new product [CP] than the base?") on seven-point scales ranging from –3 ("much less") to +3 ("much more") (r = .80). Finally, the perceived sunk cost associated with the base product was measured with a scale adopted from Okada (2006); namely, participants who owned the base product indicated the extent to which they would consider the money they already spent on their currently owned base (i.e., a PDA or an MP3 player) if they were to upgrade to this CP (on a seven-point scale from "not be a consideration at all" to "an extremely significant consideration").

Results

Manipulation check. Consistent with the pretest, the PDA was associated with relatively more utilitarian than hedonic value (5.69 versus 4.93; $t_{(998)} = 20.60$, $p = .000$), whereas the opposite was true for the MP3 player (4.91 versus 5.73; $t_{(999)} = -22.67$, $p = .000$).[15] In addition, for the added functionalities, the three manipulation criteria were met; namely, (1) for utilitarian versus hedonic value, the Yellow Pages was associated with relatively more utilitarian than hedonic value (5.30 versus 4.23; $t_{(998)} = 26.00$, $p = .000$), whereas the opposite was true for satellite radio (5.07 versus 5.66; $t_{(999)} = -16.90$, $p = .000$); (2) for goal congruence, the Yellow Pages was congruent and satellite radio was incongruent for PDA (4.43 versus 3.33; $t_{(997)} = 10.98$, $p = .000$), whereas for the MP3 player, the Yellow Pages was incongruent and satellite radio was congruent (2.60 versus 4.52; $t_{(998)} = -19.44$, $p = .000$); and (3) for overall perceived value, the Yellow Pages and satellite radio did not differ in their overall perceived value (5.77 versus 5.81; $t_{(1997)} = -.35$, $p = .725$).

The effects for prior ownership were based on the premise that owners valued the base product more than nonowners. As expected, owners had a much greater overall perceived value for the base than nonowners for both the PDA (7.73 versus 6.62; $t_{(997)} = 7.43$, $p = .000$) and the MP3 player (7.17 versus 5.72; $t_{(998)} = 11.28$, $p = .000$).

Effect of goal congruence on the incremental value of CPs. Table 2 shows the results for the incremental value of the four CPs for both OIV and IWTP. It was proposed that goal congruence would have an asymmetric effect on the evaluation of CPs with a utilitarian versus hedonic base product. As predicted, there was a significant interaction between the two factors of goal congruence and the type of base product (utilitarian versus hedonic) on both the mea-

---

[14] Respondents were provided with a range of reference prices for the two base products. Specifically, it was stated that, "on average, a PDA (an MP3 player) usually costs anywhere from $150 to $300 ($200 to $400) in the market."

[15] According to these ratings, it appears that the two base products—the PDA and the MP3 player—are not strongly discriminated in terms of their utilitarian versus hedonic value (the ratings seem close). However, this should not be an issue in the current context, because these close ratings provide a stronger test for the proposed hypotheses.

TABLE 2
The Incremental Value of CPs with (Congruent/Incongruent) New Functionalities Added to a Utilitarian Versus Hedonic Base

	Utilitarian Base Product (PDA)				Hedonic Base Product (MP3 Player)			
	Congruent, Utilitarian Functionality (YEL-PDA)	Incongruent, Hedonic Functionality (SAT-PDA)	t (p Value)[a]	γ^b	Congruent, Hedonic Functionality (SAT-MP3)	Incongruent, Utilitarian Functionality (YEL-MP3)	t (p Value)	γ
OIV	1.26	1.45	−1.71 (044)	.12	1.83	.49	11.21 (.000)	.71
IWTP	$55.38	$73.45	−3.63 (.000)	.23	$84.26	$47.43	6.66 (.000)	.43
Incremental pleasure (hedonic value)	.64	1.80	−10.39 (.000)	.66	1.84	−.07	16.10 (.000)	1.02
N	499	500			500	500		

[a] All p values are based on a one-tailed test of significance.
[b] The gammas (γ) are effect sizes that are calculated as the difference in means between the congruent and incongruent added functionalities divided by the pooled standard deviations (Cohen 1988).

sures of incremental value (for OIV, $F_{(1,\ 1995)} = 88.74$, $p = .000$; for IWTP, $F_{(1,\ 1959)} = 54.43$, $p = .000$). Further analyses were performed to test the asymmetric additivity effect and the incremental value of CPs with a utilitarian versus hedonic base product.

Asymmetric additivity effect. As H_1 predicted, the incremental value of adding a hedonic functionality to a utilitarian base (i.e., for SAT-PDA) was significantly greater than that of adding an equally valued utilitarian functionality to a hedonic base product (i.e., for YEL-MP3) for both OIV (1.45 versus .49; $t_{(998)} = 7.81$, $p = .000$) and IWTP ($73.45 versus $47.43; $t_{(984)} = 5.23$, $p = .000$). It was proposed that the asymmetric effect was due to the enhancing versus diluting effect of the incongruent functionalities on the incremental pleasure of using utilitarian versus hedonic base products, respectively. As Table 2 shows, the effect on incremental pleasure was supported. Specifically, the utilitarian base product became more pleasurable to use when an incongruent, hedonic functionality was added, whereas the hedonic base product became less (or not more) pleasurable to use when an incongruent, utilitarian functionality was added (1.80 versus –.07; $t_{(998)} = 14.99$, $p = .000$).

CPs with a utilitarian base. It was proposed that for CPs with a utilitarian base, adding a congruent functionality (e.g., YEL-PDA) would be subject to diminishing returns,[16] whereas adding an incongruent functionality (e.g., SAT-PDA) would enhance the hedonic value of the base. Accordingly, CPs with an incongruent addition would have more incremental value than those with a congruent one (H_{2a}). As Figure 2 shows, H_{2a} was supported for both the measures of incremental value; specifically, (1) OIV was greater for SAT-PDA than for YEL-PDA (1.45 versus 1.26; $t_{(997)} = 1.71$, $p = .044$), and (2) IWTP was higher for SAT-PDA than for YEL-PDA ($73.45 versus $55.38; $t_{(979)} = 3.63$, $p = .000$).

It was proposed that this effect was due to the enhancing effect of the incongruent functionality on the incremental pleasure (i.e., hedonic value) of using the utilitarian base (H_{2b}). As Table 2 shows, for CPs with a utilitarian base, adding an incongruent functionality was associated with greater (enhanced) incremental pleasure (i.e., hedonic value) than the addition of a congruent functionality (1.80 versus .64 for SAT-PDA versus YEL-PDA, respectively; $t_{(998)} = 10.39$, $p = .000$), in support of H_{2b}. Further analyses using the Sobel test also supported the predicted mediation of incremental pleasure. Regression analyses revealed that for CPs with a utilitarian base, goal congruence (congruent = 0, and incongruent = 1) had a significant effect on incremental pleasure ($\beta = 1.16$, SE = .11; t = 10.39; $p = .000$).[17] When both goal congruence and incremental pleasure were used as predictors for the incremental value of CPs, the coefficients were as follows: For OIV, $\beta_{\text{goal congruence}} = -.70$, SE = .07; t = –9.57; $p = .000$; and $\beta_{\text{incremental pleasure}} = .77$, SE = .02; t = 38.84; $p = .000$; for IWTP, $\beta_{\text{goal congruence}} = 1.47$, SE = 4.94; t = .30; $p = .767$; and $\beta_{\text{incremental pleasure}} = 14.64$, SE = 1.33; t = 10.98; $p = .000$).[18] The Sobel test revealed a significant mediation

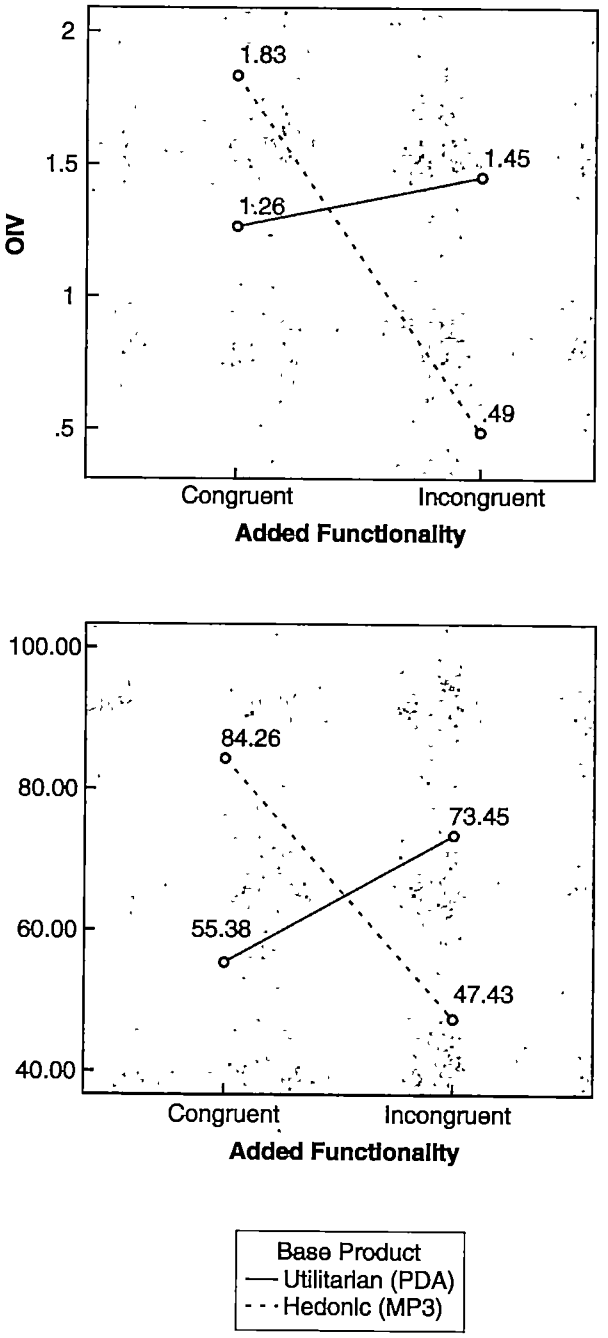
---

[16] To investigate the diminishing marginal utility of adding a congruent functionality (in YEL-PDA), OIV (i.e., the incremental value of the CP) was regressed on the perceived value of the base product. Because it was expected that OIV would increase at a diminishing rate as the base value increased, regression analyses were performed by mean splitting the perceived value of the base. The results from these analyses revealed that the OIV for YEL-PDA increased at a lower rate for the high- versus low-base-value group ($\beta = .44$ versus .49). Conversely, for the CP with an incongruent functionality (i.e., SAT-PDA), the OIV increased at a higher rate for the high- versus low-base-value group ($\beta = .49$ versus .42). These results suggest that for CPs with a utilitarian base, adding a congruent functionality was subject to diminishing returns, but adding an incongruent functionality was not.

[17] All β-values reported in the article are unstandardized coefficients.

[18] The corresponding coefficients when goal congruence was used as the sole predictor were as follows: For OIV, $\beta_{\text{goal congruence}} = .19$, SE = .11; t = 1.71; $p = 088$; for IWTP, $\beta_{\text{goal congruence}} = 18.07$, SE = 4.98; t = 3.63; $p = .000$. A comparison of these coefficients with those reported previously (when both goal congruence and incremental pleasure were used as predictors) shows that the coefficient for goal congruence was reduced

FIGURE 2
The Incremental Value of CPs with a Utilitarian Versus Hedonic Base



effect of incremental pleasure for both OIV (test statistic = 10.17, $p = .000$) and IWTP (test statistic = 7.61, $p = .000$).[19]

CPs with a hedonic base. For CPs with a hedonic base, it was proposed that adding a congruent functionality (e.g., SAT-MP3) would maintain the hedonic image of the base (albeit with diminishing returns[20]), whereas an incongruent functionality (e.g., YEL-MP3) could dilute the hedonic value of the base. Accordingly, CPs with an incongruent addition would have less incremental value than those with a congruent one (H_{3a}). As Figure 2 shows, H_{3a} was supported; specifically, (1) OIV was greater for SAT-MP3 than for YEL-MP3 (1.83 versus .49; $t_{(998)} = 11.21$, $p = .000$), and IWTP was higher for SAT-MP3 than for YEL-MP3 ($84.26 versus $47.43; $t_{(980)} = 6.66$, $p = .000$).

It was proposed that this effect was due to the diluting effect of the incongruent functionality on the incremental pleasure (i.e., hedonic value) of using the hedonic base product (H_{3b}). As Table 2 shows, for CPs with a hedonic base, adding an incongruent, utilitarian functionality was associated with a lower (diluted) incremental pleasure (i.e., hedonic value) than adding a congruent, hedonic functionality (–.07 versus 1.84 for YEL-MP3 versus SAT-MP3, respectively; $t_{(998)} = -16.10$, $p = .000$), in support of H_{3b}.[21] A mediation analysis on the CPs with a hedonic base revealed that goal congruence had a significant effect on incremental pleasure ($\beta = -1.91$, SE = .12; $t = -16.10$; $p = .000$), and when both goal congruence and incremental pleasure were used as predictors for the incremental value of CPs, the coefficients were as follows: For OIV, $\beta_{\text{goal congruence}} = .24$, SE = .08; $t = 3.19$; $p = .001$; and $\beta_{\text{incremental pleasure}} = .83$, SE = .02; $t = 45.63$; $p = .000$; for IWTP, $\beta_{\text{goal congruence}} = -4.20$, SE = 5.76; $t = -.73$; $p = .467$; and $\beta_{\text{incremental pleasure}} = 17.10$, SE = 1.37; $t = 12.50$; $p = .000$.[22] The Sobel tests revealed a significant mediation of

---

(reversed) for OIV and became nonsignificant for IWTP. This change further supports the mediating role of incremental pleasure (Baron and Kenny 1986).

[19] All Sobel test statistics reported in this article were computed using an interactive, Internet-based tool available in the work of Preacher and Leonardelli (2001).

[20] Again, the diminishing marginal utility of adding a congruent functionality (in SAT-MP3) was investigated by regressing OIV after mean splitting the perceived value of the base product. The results revealed that the OIV for SAT-MP3 increased at a lower rate for the high- versus low-base-value group ($\beta = .36$ versus .46). Conversely, for the CP with an incongruent functionality (i.e., YEL-MP3), the OIV increased at a higher rate for the high- versus low-base-value group ($\beta = .40$ versus .32). These results suggest that even for CPs with a hedonic base, adding a congruent functionality was subject to diminishing returns, but adding an incongruent functionality was not. That said, and as Figure 2 shows, the OIV of the CP with a congruent functionality (SAT-MP3) was significantly higher than that for the CP with an incongruent functionality (YEL-MP3).

[21] Note that though adding an incongruent, utilitarian functionality to the hedonic base (in YEL-MP3) would have increased the perceived utilitarian value of the base, it was predicted that the associated loss in the hedonic value (incremental pleasure) would have been weighted heavily. This is evident from comparing the incremental value of YEL-MP3 with that of YEL-PDA, which both add the same functionality but to different base products. As Table 2 shows, YEL-MP3 had a significantly lower incremental value (OIV) than YEL-PDA (.49 versus 1.26; $t(997) = -6.54$, $p = .000$; effect size = .41). This suggests a significant, negative influence of loss in hedonic value in the case of YEL-MP3. Conversely, when comparing SAT-PDA with SAT-MP3 (both of which added a hedonic functionality but to different bases), there is a much smaller, negative effect of adding an incongruent, hedonic functionality to a utilitarian base (the OIVs were 1.45 versus 1.83, respectively; $t(998) = -3.43$, $p = .001$; effect size = .22).

[22] The corresponding coefficients when goal congruence was used as the sole predictor were as follows: For OIV, $\beta_{\text{goal congruence}} = -1.35$, SE = .12; $t = -11.21$; $p = .000$; for IWTP, $\beta_{\text{goal congruence}} = -36.84$, S.E. = 5.53; $t = -6.66$; $p = .000$. A com-

incremental pleasure on the relationship between goal congruence and the evaluation of CPs with a hedonic base (for OIV, test statistic = –15.05, $p = .000$; for IWTP, test statistic = –9.82, $p = .000$).

Ownership effects in the evaluation of CPs. It was predicted that prior ownership of the base product would have a moderating effect on the evaluation of CPs. Specifically, it was proposed that the effect of goal congruence on CP evaluation would be stronger for owners than for nonowners of a hedonic base product (H_{4b}), but no ownership effects were expected for CPs with a utilitarian base product (H_{4a}). An analysis of variance revealed a significant three-way interaction among the type of base product, goal congruence, and prior ownership of the base on IWTP ($F_{(1, 1955)} = 11.32$, $p = .001$). However, this interaction was not significant for OIV ($F_{(1, 1991)} = .98$, $p = .324$). Further analyses revealed that for the CPs with a utilitarian base (the PDA), the interaction between goal congruence and ownership was not significant for OIV ($F_{(1, 995)} = .04$, $p = .837$), but it was marginally significant for IWTP ($F_{(1, 977)} = 3.42$, $p = .065$); for CPs with a hedonic base (the MP3 player), it was marginally significant for OIV ($F_{(1, 996)} = 2.87$, $p = .091$) and highly significant for IWTP ($F_{(1, 978)} = 8.95$, $p = .003$). The mean incremental values (both OIV and IWTP) associated with the four CPs appear in Table 3 separately for owners versus nonowners of the corresponding base products. As Table 3 shows, in the case of a utilitarian base (the PDA), the effect sizes between owners and nonowners did not differ for OIV, and they were only marginally stronger for owners than for nonowners for IWTP (as per the statistical tests for interaction effects). These results support H_{4a} that the effect of goal congruence is equal for owners and nonowners for CPs with a utilitarian base. Conversely, for the hedonic base product (the MP3 player), the effect sizes obtained for goal congruence were stronger for owners than for nonowners, marginally so for OIV and significantly so for IWTP (as per the statistical tests for the interaction effects). These results support H_{4b} that for CPs with a hedonic base, owners show stronger effects (on incremental value) for the goal congruence of the added functionality than nonowners.

Ownership effects in CPs due to the perceived change in image of the base product. The ownership effects in CPs were predicted to be the result of owners' greater sensitivity

---

parison of these coefficients with those reported previously (when both goal congruence and incremental pleasure were used as predictors) shows that the coefficient for goal congruence was reduced (reversed) for OIV and became nonsignificant for IWTP. This change also supports the mediating role of incremental pleasure (Baron and Kenny 1986).

TABLE 3
Prior Ownership Effects in the Evaluation of CPs with a Utilitarian Versus Hedonic Base

	Utilitarian Base Product (PDA)					
	Owners			**Nonowners**		
	Congruent, Utilitarian Functionality (YEL-PDA)	**Incongruent, Hedonic Functionality (SAT-PDA)**	**γ[a]**	**Congruent, Utilitarian Functionality (YEL-PDA)**	**Incongruent, Hedonic Functionality (SAT-PDA)**	**γ**
OIV	1.35	1.49	.08	1.24	1.44	.12
IWTP	$65.37	$100.24	.36	$52.31	$65.55	.19
Perceived sunk costs	4.87	4.41	.24	—	—	—
Perceived change in image[b]	1.24	.94	.23	1.15	.84	.24
N	116	114		383	386	
	Hedonic Base Product (MP3)					
	Owners			**Nonowners**		
	Congruent, Hedonic Functionality (SAT-MP3)	**Incongruent, Utilitarian Functionality (YEL-MP3)**	**γ**	**Congruent, Hedonic Functionality (SAT-MP3)**	**Incongruent, Utilitarian Functionality (YEL-MP3)**	**γ**
OIV	1.88	.32	.82	1.80	.65	.61
IWTP	$98.11	$44.00	.60	$71.81	$50.66	.26
Perceived sunk costs	4.61	4.41	.10	—	—	—
Perceived change in image[b]	1.65	–.31	1.35	1.52	.09	1.00
N	233	242		267	258	

[a] The gammas (γ) are effect sizes that are calculated as the difference in means between the congruent and incongruent added functionalities divided by the pooled standard deviations (Cohen 1988).

[b] Perceived change in image is measured as the incremental change in utilitarian image for the utilitarian base and as the incremental change in hedonic image for the hedonic base product.

(than nonowners) to the perceived change in image of the hedonic base (but not so for a utilitarian base). As Table 3 shows, for CPs with a utilitarian base product (the PDA), there was no difference in the perceived change in image of the base between owners and nonowners (i.e., the interaction between goal congruence and prior ownership was not significant; $F_{(1, 995)} = .01$, $p = .938$), in support of H_{5a}. Conversely, for CPs with a hedonic base (the MP3 player), the perceived change in image was stronger for owners than for nonowners (i.e., this interaction was significant; $F_{(1, 996)} = 8.48$, $p = .004$). Furthermore, as Table 3 shows, although owners of the hedonic base product did not perceive a significantly larger enhancement in the hedonic image than nonowners when a congruent, hedonic functionality was added (1.65 versus 1.52 for SAT-MP3; $t_{(498)} = 1.13$, $p = .259$), they perceived a significantly larger dilution effect than nonowners when an incongruent functionality was added to the base (–.31 versus .09 for YEL-MP3; $t_{(498)} = -2.84$, $p = .005$). These analyses support H_{5b} that for CPs with a hedonic base, prior owners are more sensitive than nonowners to the perceived change in hedonic image of the base. Moreover, this effect is driven by the owners' higher sensitivity than nonowners to the dilution effect of adding incongruent, utilitarian functionalities to the hedonic base (as in YEL-MP3).

The effect of perceived sunk costs was also explored to compare the findings in the current context with those of Okada (2006). As Table 3 shows, for the utilitarian base product (the PDA), owners indicated a much greater concern for perceived sunk costs when a congruent, utilitarian functionality was added than when an incongruent, hedonic one was added (4.87 versus 4.41 for YEL-PDA versus SAT-PDA; $t_{(224)} = 1.82$, $p = .036$). However, the Sobel test revealed that the mediation of perceived sunk costs was not significant for OIV (test statistic = –1.36, $p = .174$) or for IWTP (test statistic = –1.42, $p = .156$). In the case of owners of the hedonic base, there was no difference in the perceived sunk costs for CPs with a congruent versus incongruent functionality (4.61 versus 4.41 for SAT-MP3 versus YEL-MP3; $t_{(470)} = 1.11$, $p = .267$). These results support the previous claim that perceived sunk costs are not relevant here, because the current context of CPs is focused solely on nonalignable enhancements (i.e., those achieved by adding new functionalities to a base).

Discussion

This experimental study (conducted with a representative sample of the target market) supported the theoretical model in Figure 1. Specifically, it was shown that in the case of CPs with a utilitarian base, adding an incongruent, hedonic functionality was valued more than adding a congruent, utilitarian one (SAT-PDA had a greater OIV and IWTP than YEL-PDA), whereas for CPs with a hedonic base product, adding an incongruent, utilitarian functionality was valued less than a congruent, hedonic one (YEL-MP3 had a lower OIV and IWTP than SAT-MP3). The perceived incremental pleasure (hedonic value) in using the base product mediated the latter asymmetry. That is, whereas an incongruent, hedonic addition enhanced the pleasure of using a utilitarian base (e.g., in SAT-PDA), an incongruent, utilitarian addition diluted the hedonic image of the hedonic base (e.g., in YEL-MP3). That said, the effect sizes were small in the case of the utilitarian base (the PDA), and they were medium to large for the hedonic base (the MP3 player) (see Table 2). The small effect in the case of the PDA may have been because it was not strongly discriminated on its utilitarian versus hedonic associations (5.69 versus 4.93, respectively). This may have prevented a large enough contrast effect when an incongruent, hedonic functionality was added to the base. Although the lack of discrimination provided a stronger test for the proposed hypotheses, it could have resulted in a smaller effect size. Further research with a relatively pure utilitarian base could address this issue.

The moderating role of prior ownership of the base on the evaluation of CPs was also explored. Ownership effects occurred only in CPs with a hedonic base product, not for those with a utilitarian base. Given the owners' greater attachment and need to maintain the image for a hedonic base product, the moderating effect of prior ownership was stronger for a CP with a hedonic base than for a utilitarian one. Furthermore, the perceived change in image of the base mediated this ownership effect, such that owners showed a larger dilution effect than nonowners when an incongruent, utilitarian functionality was added to a hedonic base (i.e., in YEL-MP3). That being said, perceived sunk costs were not relevant in the evaluation of CPs by prior owners of the base.

General Discussion

Theoretical Implications

Convergence in the consumer electronics, computers, and communications industries has allowed the introduction of CPs with new (and disparate) functionalities added to existing base products (e.g., mobile television on a cell phone, Internet on a handheld gaming device). Much of the existing literature has investigated the addition of new features that are related to the base products' consumption goals (see Meyers-Levy and Tybout 1989; Nowlis and Simonson 1996). However, the addition of new functionalities with incongruent and often unrelated consumption goals has not been addressed (in the context of combining two concepts/products, cf. Gill and Dube 2007; in the context of using comparisons with existing functionalities, cf. Ziamou and Ratneshwar 2003). Furthermore, the role of utilitarian versus hedonic goals (and value) associated with the added functionality and the base product of a CP has not been examined. Using these alternative goal associations, the current research proposes a theoretical model for the evaluation of CPs that have congruent versus incongruent functionalities added to a utilitarian versus hedonic base. As Figure 1 shows, the key aspect of this model is the asymmetry in the evaluation of CPs with incongruent additions to a utilitarian versus hedonic base (i.e., the asymmetric additivity effect). On the basis of this asymmetry, it is proposed and found that CPs with a utilitarian base gain more from an incongruent addition than from a congruent one; however, the opposite is true for CPs with a hedonic base; they gain less from an incongruent addition than from a congruent one. This proposed theoretical model extends the existing approaches on new product perception to the context of CPs, and it adds to the sparse literature on the addition of new functionalities to products.

The findings on the prior ownership effects extend Okada's (2006) research in the context of upgraded (enhanced) new products. Okada compares alignable enhancements (in existing features) with nonalignable ones (by adding new features) and finds that owners value the latter more than the former because of the lower perceived sunk costs. The current research examines solely nonalignable enhancements in the base product of CPs (by adding new functionalities), and in this context, the perceived sunk costs were not relevant. Moreover, although the incongruent functionalities were valued more (than congruent ones) by owners of a utilitarian base, the opposite was true for owners of a hedonic base product. These findings provide a boundary condition to Okada's work; specifically, prior owners do not always positively value nonalignable enhancements (upgrades). Furthermore, given that ownership effects were observed only for hedonic base products and not utilitarian ones, this specifies another boundary condition in which owners' upgrade decisions may be similar to nonowners' new purchases.

Finally, the current research has relevance to the recent work on "feature fatigue" (i.e., the detrimental effect of loading products with numerous features; Rust, Thompson, and Hamilton 2006; Thompson, Hamilton, and Rust 2005).[23] That research finds that though adding more features (or functionalities) increases the perceived capability (i.e., the number of benefits provided) of the feature-loaded product, it simultaneously reduces its perceived usability (i.e., the ease of use). This reduction in usability is felt even more after the product is used (experienced utility), which could eventually have a detrimental effect on the long-term satisfaction with the feature-loaded product. Rust, Thompson, and Hamilton (2006) recommend that firms should aim to maintain a balance between the capability and the usability of a product with added features (functionalities) to ensure long-term profitability. First, the current findings suggest that the goal congruence of the added functionality moderates the perceived capability of the CP (e.g., SAT-PDA is more capable than YEL-PDA) and that the moderation is asymmetric for a utilitarian versus hedonic base (e.g., YEL-MP3 is less capable than YEL-PDA). In other words, not all feature additions (or new functionalities) will increase perceived capability equally. Second, to the extent that hedonic functionalities are easier to imagine (MacInnis and Price 1987), they might also increase the perceived usability of CPs. As a result, the perceived feature fatigue may be lower for hedonic additions than for utilitarian additions in CPs. That said, it would be relevant to investigate whether this also translates into higher experienced utility (after use) for hedonic functionalities than for utilitarian ones. In summary, although feature fatigue would be prevalent in the case of CPs (especially when the base is already loaded with functionalities), the effect would be moderated by the goal congruence of the added functionality, and it may be felt to a lesser extent for hedonic additions than for utilitarian ones.

Managerial Implications

Owing to convergence, marketers in the consumer electronics, computers, and communications industry (e.g., the manufacturers of cell phones, televisions, PCs, MP3 players, and PDAs) have access to a potent means to upgrade their existing products. Adding new functionalities from different categories has become a widely used (and often fashionable) means to innovate and launch new CPs. The current research provides marketers a theoretical perspective to understand the incremental value of these added functionalities. In light of the goal congruence of the added functionality and the utilitarian versus hedonic image of the base, marketers can appreciate what type of added functionalities would be valued in a CP. Some recent examples highlight the applicability of the present research findings. For example, in the context of cell phones (a relatively utilitarian base product), a recent survey found that the addition of cameras (a relatively hedonic, incongruent functionality) has gained more popularity (in terms of usage) than the addition of e-mail/instant messaging (a relatively utilitarian, congruent functionality) (*CIO Insight* 2006). Similarly, in the context of hedonic products, whereas video iPods (a hedonic addition to a hedonic base product) were an instant success, the concept of Web television (i.e., Internet access on a television)—a utilitarian addition to a hedonic base—has not been successful, despite many attempts by a reputable manufacturer (Ries and Ries 2005). These examples, though not strictly controlled in terms of the relative hedonic and utilitarian value of the base and the added functionality (and the associated goal congruence or perceived value), are anecdotally consistent with the proposed theoretical model.

The ownership effects found in the context of CPs have implications for targeting these new products to specific segments. On the basis of the findings, it can be said that for CPs with a utilitarian base, both segments (prior owners and nonowners) value incongruent, hedonic enhancements more than congruent, utilitarian ones. However, for CPs with a hedonic base, prior owners are more sensitive than nonowners to the changes in the hedonic image of the base. This is particularly important in the context of existing hedonic products (e.g., iPods and handheld games), whose current owners may be particularly resistant to any additions of utilitarian functionalities. Moreover, any such additions may actually harm (dilute) the existing image (value) of the base products among their prior owners.

Limitations and Further Research

Although the current research provides evidence for the asymmetric effect in the evaluation of CPs, further investigation is needed for it to be generalizable across contexts. Two avenues of research can address these limitations. First, the current research investigated only the evaluation of CPs in the context of handheld consumer electronics products (a PDA and an MP3 player). Furthermore, these base products were not strongly discriminated on their utilitarian versus hedonic associations (see n. 15), which, for the PDAs, could have led to the small effect size.[24] Although this result is not a threat to the internal validity of the proposed theoretical model, further extension in other product categories that are purely utilitarian (e.g., printer, fax) versus purely hedonic (e.g., handheld games) would add to the robustness of the current findings. The model could also be tested in contexts other than the electronics/communications industry to investigate the generalizability of the proposed asymmetric additivity effect. For example, in the context of services, adding more utilitarian, congruent functionalities to a relatively utilitarian base (e.g., offering self-defense lessons at a fitness center) may be valued less than a hedonic addition (e.g., offering dance lessons at a fitness center), assuming they are equally valued. Conversely, for a relatively hedonic base, adding more hedonic elements (e.g., adding music kiosks in a movie theater complex) may be valued more than equally valued utilitarian additions (e.g., adding Internet access kiosks in a movie theater complex), which may dilute the image of the base. However, there would be potential moderators to these effects. For example, adding hedonic functionalities may not always enhance a utilitarian base product/service (e.g., adding music kiosks in a bank may interfere with the base service performance and reduce the perceived credibility of the provider). Similarly, in the context of hedonic base products, adding a utilitarian element may sometimes be valued (e.g., adding vitamins to an ice cream)—for example, when multiple goals are salient (e.g., both health and pleasure) and people are willing to compromise on the dilution of the hedonic value of the base (e.g., the taste of the ice cream). Extending the current work to other consumption contexts and the identification of boundary conditions/potential moderators to the proposed effects would be a useful avenue for further research.

---

[23] Note that this research on "feature fatigue" uses feature addition as synonymous with the addition of new functionalities.

A second limitation is that the current definition of goal congruence is somewhat restrictive and should be expanded to include other types of CPs. As noted previously, goal congruence has been defined in terms of congruity of both the consumption goals (benefits) and the value associated with the base product and the added functionality. This excludes CPs with added functionalities that may be incongruent in terms of their utilitarian versus hedonic value but nevertheless can enhance the existing consumption goals (benefits) of the base. For example, adding the ability to download music wirelessly (through the Internet or in a music store) to an MP3 player would entail adding an incongruent, utilitarian functionality to a hedonic base (in terms of "value"), but the benefits of the added functionality are complementary with the base (because they enhance the base goal of listening to music). Such incongruent but complementary additions to a hedonic base would certainly be valued more than additions that are noncomplementary (unrelated), as was found in the case of YEL-MP3. Such goal complementarity was also crucial in the success of adding cameras to cell phones, which not only enhanced the utilitarian base by adding an incongruent, hedonic functionality but also complemented the existing goals of the base by allowing communication through pictures (e.g., sending pictures instantly).[25] Conversely, adding the ability to play MP3 files to cell phones has been relatively less successful because of the lack of complementary goals, even though the added functionality enhanced the hedonic value of the base. These effects of goal complementarity should be investigated in further research, which would add to the understanding of which new functionalities add more value in CPs.

Conclusion

The current research investigates the evaluation of CPs formed by adding new functionalities to existing base products. It is shown that the incremental value of CPs is subject to an asymmetric additivity effect (i.e., incongruent, hedonic additions are valued more than congruent ones in a utilitarian base, whereas incongruent, utilitarian additions are valued less than congruent ones in a hedonic base). Furthermore, this effect is stronger among owners than among nonowners of the hedonic base. These findings suggest that managers should consider the goal congruence between the added functionality and the base product when developing new CPs. Moreover, the potential dilution effect of adding utilitarian functionalities to a hedonic base should be a concern, especially among prior owners of the base. Further research on the generalizability of the current findings and on the potential moderators of the proposed asymmetric effect in CPs is suggested.

---

[24] Thanks to an anonymous *JM* reviewer for pointing out this issue.

[25] In the current context, goal complementarity is defined such that an added functionality has complementary goals if it enhances the existing benefits (goals) of the base product, even though it may be incongruent in terms of the associated (hedonic or utilitarian) value. For example, adding a camera to a cell phone offers incongruent value (hedonic + utilitarian), but it complements the existing goals of the base by allowing communication through sending pictures. Conversely, an added functionality has noncomplementary goals if it is unrelated to the goals of the base (e.g., YEL-MP3 [utilitarian + hedonic]). Note that this definition is different from the traditional economic definition of complements, which refers to products that are used together to fulfill a particular consumption goal (e.g., toothpaste and toothbrush are complements because they are used together for cleaning teeth).

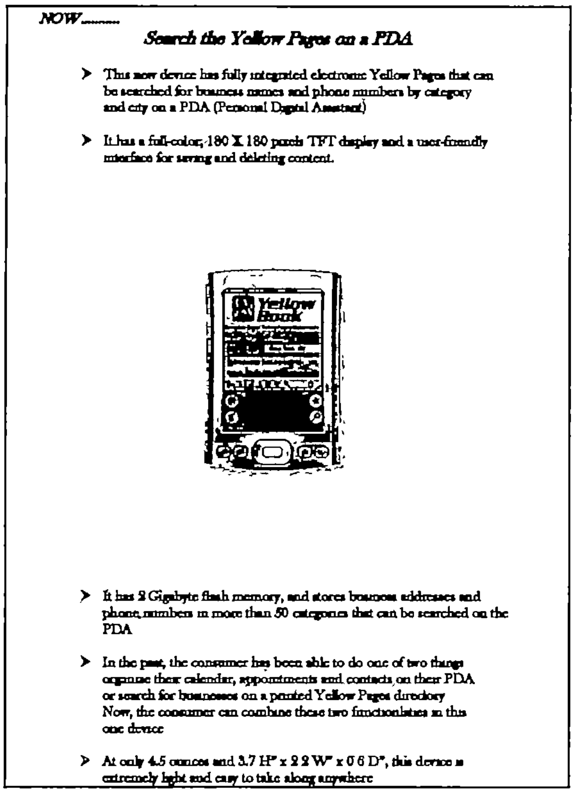
APPENDIX A
Demographic Distribution (in %) of the Survey Samples and the U.S. Population

Demographic Segment	Survey Samples: Survey 1 (YEL-PDA)	Survey Samples: Survey 2 (SAT-PDA)	Survey Samples: Survey 3 (SAT-MP3)	Survey Samples: Survey 4 (YEL-MP3)	U.S. Population
Gender					
Male	55.1	52.8	51.8	46.0	49.1
Female	44.9	47.2	48.2	54.0	50.9
Age					
18–24	19.0	17.4	12.8	18.2	13.0
25–34	24.4	25.6	25.8	26.8	19.0
35–44	33.7	28.8	29.4	28.0	21.6
45–54	22.8	27.8	30.6	25.2	18.0
55 and older	.0	.4	1.4	1.8	28.3
Income					
Under $25,000	21.2	19.8	19.4	15.6	23.9[a]
$25,000–$49,999	29.1	30.4	33.0	33.0	34.1[b]
$50,000–$74,999	25.1	24.6	22.8	24.4	19.5
$75,000 and more	24.6	25.2	24.8	27.0	22.5
Education					
Some or no high school	2.0	1.2	1.0	1.4	20.3
High school diploma	14.6	17.4	18.8	17.6	28.6
Some college	41.1	44.2	38.0	36.6	23.3
College diploma or more	42.3	37.2	42.2	44.4	27.8

[a] The percentage of the U.S. population in this group was estimated by taking five-sixths of 28.7% "under $30,000," according to the date provided by the market research firm

[b] The percentage of the U.S. population in this group was estimated by adding one-sixth of 28.7% "under $30,000" to 29.3% from "$30,000–$49,000."

APPENDIX B
Concept Description for Two of the CPs (YEL-PDA and SAT-MP3) Used in the Main Study



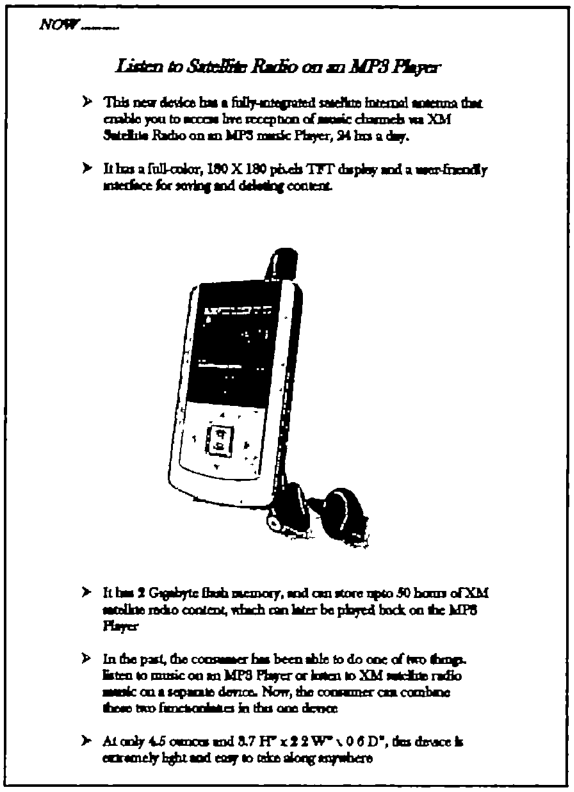
NOW……

Search the Yellow Pages on a PDA

- This new device has fully integrated electronic Yellow Pages that can be searched for business names and phone numbers by category and city on a PDA (Personal Digital Assistant)
- It has a full-color, 180 X 180 pixels TFT display and a user-friendly interface for saving and deleting content.

Yellow Book

- It has 2 Gigabyte flash memory, and stores business addresses and phone numbers in more than 50 categories that can be searched on the PDA
- In the past, the consumer has been able to do one of two things organize their calendar, appointments and contacts on their PDA or search for businesses on a printed Yellow Pages directory Now, the consumer can combine these two functionalities in this one device
- At only 4.5 ounces and 3.7 H" x 2.2 W" x 0.6 D", this device is extremely light and easy to take along anywhere



NOW……

Listen to Satellite Radio on an MP3 Player

- This new device has a fully-integrated satellite internal antenna that enable you to access live reception of music channels via XM Satellite Radio on an MP3 music Player, 24 hrs a day.
- It has a full-color, 180 X 180 pixels TFT display and a user-friendly interface for saving and deleting content.
- It has 2 Gigabyte flash memory, and can store upto 50 hours of XM satellite radio content, which can later be played back on the MP3 Player
- In the past, the consumer has been able to do one of two things. listen to music on an MP3 Player or listen to XM satellite radio music on a separate device. Now, the consumer can combine these two functionalities in this one device
- At only 4.5 ounces and 3.7 H" x 2.2 W" x 0.6 D", this device is extremely light and easy to take along anywhere

REFERENCES

Baron, Reuben M. and David Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Batra, Rajeev and Olli T. Ahtola (1990), "Measuring the Hedonic and Utilitarian Sources of Consumer Attitudes," *Marketing Letters*, 2 (2), 159–70.

Beggan, James K. (1992), "On the Social Nature of Nonsocial Perception: The Mere Ownership Effect," *Journal of Personality and Social Psychology*, 62 (2), 229–37.

Belk, Russell W. (1988), "Possessions and the Extended Self," *Journal of Consumer Research*, 15 (September), 139–68.

Carpenter, Gregory S. and Kent Nakamoto (1989), "Consumer Preference Formation and Pioneering Advantage," *Journal of Marketing Research*, 36 (August), 285–98.

CIO Insight (2006), "How Do You Use Your Phone?" 10, 26.

Cohen, Jacob (1988), *Statistical Power Analysis for the Behavioral Sciences*, 2d ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dhar, Ravi and Klaus Wertenbroch (2000), "Consumer Choice Between Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 37 (February), 60–71.

Gatignon, Hubert and Jean-Marc Xuereb (1997), "Strategic Orientation of the Firm and New Product Performance," *Journal of Marketing Research*, 34 (February), 77–90.

Gill, Tripat and Laurette Dube (2007), "What Is a Leather Iron or a Bird Phone? Using Conceptual Combinations to Generate and Understand New Product Concepts," *Journal of Consumer Psychology*, 17 (3), 202–217.

Herr, Paul M., Steven J. Sherman, and Russell H. Fazio (1983), "On the Consequences of Priming: Assimilation and Contrast Effects," *Journal of Experimental Social Psychology*, 19 (July), 323–40.

Hirschman, Elizabeth C. and Morris B. Holbrook (1982), "Hedonic Consumption: Emerging Concepts, Methods, and Propositions," *Journal of Marketing*, 46 (Summer), 92–101.

Huffman, Cynthia and Micheal J. Houston (1993), "Goal-Oriented Experiences and the Development of Knowledge," *Journal of Consumer Research*, 20 (September), 190–208.

Keller, Punam Anand and Ann L. McGill (1994), "Differences in the Relative Influence of Product Attributes Under Alternative Processing Conditions: Attribute Importance Versus Attribute Ease of Imagability," *Journal of Consumer Psychology*, 3 (1), 29–49.

Kirmani, Amna, Sanjay Sood, and Sheri Bridges (1999), "The Ownership Effect in Consumer Responses to Brand Line Stretches," *Journal of Marketing*, 63 (January), 88–101.

MacInnis, Deborah J. and Linda L. Price (1987), "The Role of Imagery in Information Processing: Review and Extensions," *Journal of Consumer Research*, 13 (March), 473–91.

Markman, Arthur B. and Douglas L. Gentner (1993), "Splitting the Differences: A Structural Alignment View of Similarity," *Journal of Memory and Language*, 32 (4), 517–35.

Martin, Ingrid M. and David W. Stewart (2001), "The Differential Impact of Goal Congruency on Attitudes, Intentions, and the Transfer of Brand Equity," *Journal of Marketing Research*, 38 (November), 471–84.

Meyers-Levy, Joan and Alice M. Tybout (1989), "Schema Congruity as a Basis for Product Evaluations," *Journal of Consumer Research*, 16 (June), 39–54.

Mitchell, Andrew A. and Peter A. Dacin (1996), "The Assessment of Alternative Measures of Consumer Expertise," *Journal of Consumer Research*, 23 (December), 219–39.

Nowlis, Stephen M. and Itamar Simonson (1996), "The Effect of New Product Features on Brand Choice," *Journal of Marketing Research*, 33 (February), 36–47.

Okada, Erica Mina (2006), "Upgrades and New Purchases," *Journal of Marketing*, 70 (October), 92–102.

Pine, B. Joseph, II, and James H. Gilmore (1998), "Welcome to the Experience Economy," *Harvard Business Review*, 76 (July–August), 97–105.

Preacher, Kristopher J. and Geoffrey J. Leonardelli (2001), "Calculation for the Sobel Test: An Interactive Calculation Tool for Mediation Tests," (accessed April–August 2007), [available at http://www.psych.ku.edu/preacher/sobel/sobel.htm].

Ratneshwar, S., Allan D. Shocker, June Cotte, and Rajendra K. Srivastava (1999), "Product, Person, and Purpose: Putting the Consumer Back into Theories of Dynamic Market Behavior," *Journal of Strategic Marketing*, 7 (3), 191–208.

Ries, Al and Laura Ries (2005), *The Origin of Brands: Discover the Natural Laws of Product Innovation and Business Survival.* New York: HarperCollins.

Rust, Roland T., Debora Viana Thompson, and Rebecca W. Hamilton (2006), "Defeating Feature Fatigue," *Harvard Business Review*, 84 (February), 98–107.

Schwarz, Norbert and Herbert Bless (1992), "Constructing Reality and Its Alternatives: Assimilation and Contrast Effects in Social Judgment," in *The Construction of Social Judgment*, L.L. Martin and A. Tesser, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 217–45.

Shiv, Baba and Alexander Fedorikhin (1999), "Heart and Mind in Conflict: The Interplay of Affect and Cognition in Consumer Decision Making," *Journal of Consumer Research*, 26 (December), 278–92.

Stevens, Stanley Smith (1986), *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects*. Oxford: Transaction Books.

Thompson, Debora Viana, Rebecca W. Hamilton, and Roland T. Rust (2005), "Feature Fatigue: When Product Capabilities Become Too Much of a Good Thing," *Journal of Marketing Research*, 42 (November), 431–42.

Tversky, Amos and Daniel Kahneman (1991), "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics*, 106 (4), 1039–1061.

Voss, Kevin E., Eric R. Spangenberg, and Bianca Grohmann (2003), "Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude," *Journal of Marketing Research*, 40 (August), 310–40.

Yoffie, David A. (1997), *Competing in the Age of Digital Convergence*. Boston: Harvard Business School Press.

Ziamou, Paschalina (Lilia) and S. Ratneshwar (2003), "Innovations in Product Functionality: When and Why Are Explicit Comparisons Effective," *Journal of Marketing*, 67 (April), 49–91.

Richard G. McFarland, James M. Bloodgood, & Janice M. Payan

Supply Chain Contagion

Drawing on research from the interfirm relationship, marketing channels, operations management, and network theory literature and on the basis of qualitative depth interviews, the authors identify a new phenomenon they call "supply chain contagion." Supply chain contagion is the propagation of interfirm behaviors from one dyadic relationship to an adjacent dyadic relationship within the supply chain. Contagion can occur inadvertently and with or without the knowledge of the affected parties. Using institutional theory, the authors develop a conceptual model that predicts the conditions under which contagion is likely to occur. Although contagion may take the form of any number of interfirm behaviors, operationally, the authors focus on whether the downstream influence strategies that manufacturers use with their dealers are imitated by these same dealers with end customers. They conduct conclusive research using a sample of 151 vertically linked manufacturer–dealer–customer supply chain triads and explain a large variance ($R^2 = .30$) in the use of downstream influence strategies in terms of supply chain contagion. Given extensive prior empirical support for alternative antecedents to influence strategy usage, the explanatory power of supply chain contagion is impressive and suggests that how intermediaries treat end customers is explained, to a large degree, as the intermediaries simply imitating how their suppliers treated them. In addition to the identification of a new theoretical concept, the study provides empirical support for the effects of both macro- and microinstitutional factors on interfirm behavior. Specific factors that are positively related to the level of manifest contagion are environmental uncertainty and the perceived similarity and frequency of contact between boundary personnel; dependence asymmetry has a negative effect on manifest contagion. Managers and boundary-spanning personnel who are aware of supply chain contagion effects should be better able to influence strategically the behavior of channel partners and may be better inoculated against their own unintended imitation of other organizations within their supply chain.

Keywords: supply chain, institutional theory, network theory, operations management, influence strategies

Research on interfirm governance and supply chain management highlights the importance of interpersonal relationships and factors such as trust, cooperation, and relational norms (Granovetter 1985; Heide and John 1992; Lee, So, and Tang 2000). Network theorists suggest that the entire network of organizations within which firms and individuals are embedded is a social network (e.g., Grewal, Lilien, and Mallapragada 2006; Ibarra, Kilduff, and Tsai 2005). Firms in these networks of relationships provide examples of behavior that is often imitated by other network members (Henisz and Delios 2001). Institutional theory proposes that extensive noneconomic motivations exist that strongly shape the form and behavior of firms (Dacin, Oliver, and Roy 2007). Together, this implies that a strong motivating force behind firm behavior is socially based and that it is embedded within personal relationships, institutions, and interconnected organizational relationships (Anderson, Håkansson, and Johanson 1994; Moran 2005).

As Grewal and Dharwadkar (2002, p. 82) note, however, "researchers have largely overlooked the ubiquitous influence of the institutional environment and how interorganizational relationships such as marketing channels are embedded in [the] larger social context." Wathne and Heide (2004) note that prior research has focused predominantly on individual dyads when developing theory for business-to-business relationships, with a ceteris paribus assumption for all other relationships. In expanding our view of the supply chain beyond dyadic interactions and explicitly incorporating social contextual motivations for firm behavior, we identify a phenomenon we call "supply chain contagion." Supply chain contagion is the propagation of interfirm behaviors from one dyadic relationship to an adjacent dyadic relationship within the supply chain.

Supply chain contagion can occur without the knowledge of affected firms. This may be particularly problematic when firms imitate behaviors that are inappropriate or even detrimental to their relationships with other firms. In the exploratory stage of this study, we found that contagion is common. However, it appears that supply chain members are acting on the assumption that the effects of their actions with each other are confined to the dyad and do not spill over into other relationships. Thus, for example, manufacturers and suppliers may unwittingly be influencing the downstream behaviors of organizations that handle and market their products in ways that are detrimental to achieving their goals.

Richard G. McFarland is Associate Professor and L.L. McAninch Chair of Business Administration, Department of Marketing (e-mail: mcfarlan@ksu.edu), and James M. Bloodgood is an associate professor, Department of Management (e-mail: jblood@ksu.edu), Kansas State University. Janice M. Payan is Associate Professor of Marketing, Monfort College of Business, University of Northern Colorado (e-mail: janice.payan@unco.edu). The authors thank Michael Ahearne, Ravi Dharwadkar, Ashwin Joshi, Eric Higgins, James Higgins, Mark Pagell, James Reardon, Raymond Rody, and Yun-Oh Whang for their help. They are also grateful to the anonymous *JM* reviewers for their helpful guidance and insights. This article was published while the first author was on sabbatical leave and a visiting professor at Yonsei University in Seoul, Korea.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

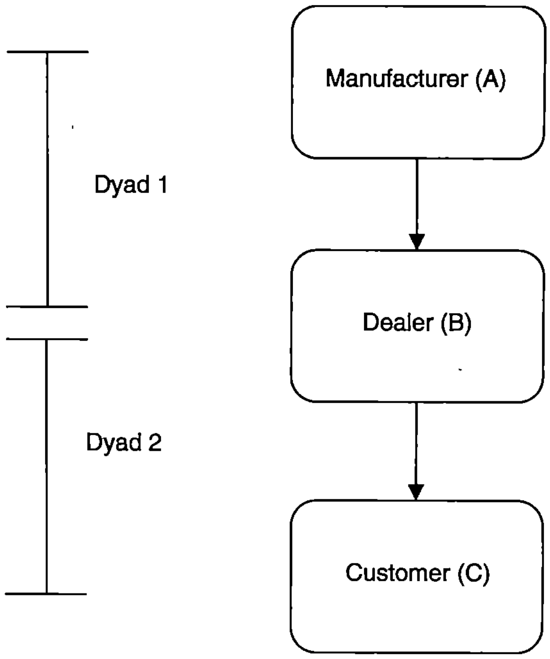
Our theoretical framework, which predicts the conditions under which supply chain contagion is likely to occur, is guided by institutional theory. Grewal and Dharwadkar (2002) develop a rich theoretical model that highlights the important role of institutions in both the structure and the behavior of organizations. They demonstrate the linkage between institutional theory and the political economy in explaining firm-level behavior within dyadic channel relationships. Specifically, their model develops propositions that examine how institutional pressures "facilitate or hinder transactions that occur within [channel] dyads" (p. 89). Although institutional theory originates primarily from a macroenvironmental perspective (e.g., national culture, regulatory institutions), we argue that it is important to consider both macro- and microinstitutional pressures in explaining convergence in interfirm behaviors within the supply chain.

Because contagion effects can have both positive and negative outcomes in the supply chain, firms that are aware of these effects should be better able to control the downstream interactions between their intermediaries and end customers. Thus, the primary purpose of this article is to identify and confirm the existence of supply chain contagion and to examine its theoretical basis through the lens of institutional theory. Operationally, we focus on whether the downstream influence strategies that manufacturers use with their dealers are imitated by these dealers with end customers. In the following sections, we discuss the background for this study and the in-depth field interviews that we initially conducted to identify the contagion phenomenon. Next, we review the literature and formulate our conceptual framework and hypotheses. This is followed by the presentation of the empirical study, including a discussion of the triadic sample and data collection, a test for measurement invariance across groups, and the results of hypotheses testing. Finally, we discuss the theoretical and managerial implications of the study and offer directions for further research.

Background

Figure 1 depicts a vertical supply chain that involves dyadic relationships at two levels. The first-level dyad involves the relationship between the manufacturer and the dealer (A → B), and the second-level dyad involves the relationship between the dealer and the end customer (B → C). We examine whether specific actions the manufacturer takes with dealers are imitated by the dealers with their end customers. In our empirical study, operationally, we focus on whether the downstream influence strategies that manufacturers use with specific dealers result in imitative influence strategy usage by the same dealers with their downstream customers. Thus, contagion is present when the influence strategies used by B with C can be partially explained as mirroring the influence strategies used by A with B. Our study uses data from the agricultural farm equipment industry, in which dealers sell both products and services to companies and consumers. The data are matched on a one-to-one-to-one basis, in which, for example, a specific dealer reports on the downstream actions of the manufacturer in the first dyad and, in turn, a customer of that same dealer reports on the dealer's downstream actions in the second dyad (i.e., non-self-reported measures).

FIGURE 1
Triadic/Dual Dyadic Supply Chain Configuration



Notes: Arrows indicate the direction of influence strategy behaviors examined in this study.

Field Interviews

Because the supply chain contagion phenomenon has not been identified previously, we initially conducted 23 in-depth field interviews with dealers and wholesalers from 14 industries in different regions of the United States to find out more about this phenomenon (see Table 1). We focused primarily on product-based wholesalers and, to a lesser extent, on retailers; however, in addition, we included several service providers to have a broadly diversified sample. We found that intermediaries frequently imitate the downstream behaviors of manufacturers/suppliers (e.g., Dyad 1 in Figure 1) in their interactions with end customers (e.g., Dyad 2 in Figure 1). This occurs even when there appears to be little economic rationale for these actions, which raises the question, "Why do firms engage in these behaviors, particularly if there is no economic benefit for doing so?"

As an example of this phenomenon, a liquor manufacturer used a noncoercive influence strategy approach when dealing with a downstream distributor, highlighting the value of its product as having an exclusive image (which would presumably provide the distributor with a positive way to distinguish itself from competitors). In turn, the distributor used this same "exclusive image" argument when dealing with its customers, even when some customers did not view this exclusivity as particularly valuable and, in some cases, even when customers considered it detrimental

TABLE 1
Characteristics of In-Depth Field Interviews

Industry	Number of Respondents in Category
Agricultural equipment dealers	2
College furniture dealer	1
Construction supply wholesaler	2
Copier machine dealer	1
Farm supply wholesaler	3
Fastener, nuts, and bolts suppliers	4
Financial services broker	1
Industrial equipment rentals	1
Industrial equipment dealer	1
Liquor and wine dealers	2
Prison supply company	1
Retail boat dealer	1
Retail furniture dealer	1
Tool supply company	2

(e.g., consumers who want to fit in rather than be different). One owner/manager in the financial services industry even coined a term for supply chain contagion: "the matching principle." She noted that financial planners often use the same strategies and messages with their clients that their broker dealers (in this case, analogous to their suppliers) used with them. In some cases, this occurred even when it appeared that the messages relevant to the financial planner were not relevant for the financial planner's customers. She could only guess why this imitation occurred. One guess was that the broker was considered credible, and the planner may have thought that he or she could obtain credibility with his or her own clients by imitating the broker's strategies and messages.

Theory

Institutional Theory

We examine the antecedents of supply chain contagion using an institutional theoretical framework. Researchers have increasingly acknowledged the importance of institutional theory in explaining firm behavior (e.g., Grewal, Comer, and Mehta 2001; Handelman and Arnold 1999; Homburg, Workman, and Krohmer 1999; Srinivasan, Lilien, and Rangaswamy 2002), whereas others have called for empirical research examining key aspects of institutional theory in marketing channels (Grewal and Dharwadkar 2002). Because firm behaviors and interactions occur within a social arena, a theoretical perspective that accounts for the impact of the social environment, rather than one that applies an economically rational perspective alone, is critical for a more complete understanding of the causes and effects of organizational behaviors within supply chain systems (Scott 1995).

People rarely start from scratch when undertaking a new course of action; rather, they are influenced by the presence of symbols and habits that exist within the environment (March and Simon 1958). People look to others to see what they have done or are doing under similar circumstances. This process leads to convergence within organizational fields and to the establishment of institutions. However, it is also important to note that these processes often do not occur consciously (March and Olsen 1984; Oliver 1991). After they are formed, institutions both empower and constrain the activities of organizations (Grewal and Dharwadkar 2002). To study these processes more deeply, it is important to distinguish between the pressures that compel convergence and the homogeneous forms or institutions that arise from these pressures (Frumkin and Kaplan 2006). In this study, the focus is on institutional pressures that lead to convergence in interfirm behaviors. Institutional pressures can arise from the presence of the regulatory environment, shared forms of cognitions, established relational norms, and so forth.

Institutional theory was first developed in the fields of sociology and political science with contributions from noted authors, including Peter Berger, Thomas Luckmann, Emile Durkheim, Karl Marx, and Max Weber (Scott 1995). Given its theoretical origins, institutional theory is often discussed and viewed from a macrosocietal level, and the management literature has primarily focused on the structure and form of firms. However, it is important to note that many authors have explicitly applied institutional theory to explain individual firm-level behaviors, firm-to-firm behaviors, and behaviors of individuals within firms (e.g., DiMaggio and Powell 1983; Jepperson 1991; Scott 1983). Grewal and Dharwadkar (2002) demonstrate the applicability of institutional theory in explaining firm-level behavior within dyadic channel relationships. Chatterjee, Grewal, and Sambamurthy (2002, p. 68) state that institutions affect technology assimilations within firms and note that institutions shape "the behaviors and cognitions of individuals within" firms. Nevertheless, much of this research, even that which uses institutional theory to explain firm-level and interfirm-level behaviors, views institutions and institutional pressures only at the macro level (e.g., national culture, regulations and laws).

Despite this, many prominent institutional theorists hold that institutions originate and function at a variety of levels (e.g., DiMaggio and Powell 1991; Jepperson 1991; Scott 1983). Thus, for example, the source of institutionalization can be an organizational field or a dyadic relationship. The *Academy of Management Journal* (in 2002, Vol. 45, Iss. 1) and *American Behaviorist Scientist* (in 2006, Vol. 49, Iss. 7) published special issues highlighting microinstitutional analysis. Henisz and Delios (2001) state that firms tend to imitate organizations with which they have social ties. Galaskiewicz and Wasserman (1989) find that though firms may mimic those they perceive as successful, it is more likely that firms will imitate organizations with which they have ties through boundary-spanning personnel. Zucker (1977) notes that macrolevel and microlevel institutions are "inextricably intertwined" and that the process of institutionalization occurs even at the interpersonal level. For this reason, in this study, we focus on both macroinstitutional pressures (e.g., environmental uncertainty) and microinstitutional pressures that can develop through the interpersonal interactions of boundary-spanning personnel (e.g., frequency of contact) and the effects of these pressures on interfirm behaviors.

Theoretical Framework

We refer to the three forms of imitation pressures or motives identified by DiMaggio and Powell (1991) as reflexive imitation, compliant imitation, and normative imitation motives.[1] Each of these three motivations suggests testable hypotheses relevant to predicting the occurrence of supply chain contagion. Reflexive imitation is a firm's standard response to environmental uncertainty. As DiMaggio and Powell (1983) note, when faced with environmental uncertainty, firms may model their behavior on that of other firms. Organizations and individuals may also mimic similar organizations or similar individuals to enhance their own legitimacy. Thus, environmental uncertainty and similarity are two primary foci of reflexive imitation. Compliant imitation is a firm's response to interorganizational dependence. We focus primarily on the degree of cohesiveness stemming from dependencies between firms as the antecedent to this type of imitation. Although firms also comply with cultural and regulatory institutions, these institutions fall outside the domain of this study. Normative imitation stems from the high degree of socialization and interaction that often occurs between members of the same organizational environment; when these members interact, they reinforce and spread norms of behavior among themselves (Scott 1983).

Our theory holds that contagious behaviors can take many forms; thus, although several other variables could potentially have been used, we concentrate on influence strategies as the focal behaviors that may be imitated. Influence strategies are defined as the content of what is communicated between one channel member and another in an attempt to influence that firm's behaviors to achieve business objectives (Boyle et al. 1992; Payan and McFarland 2005). They are the communications through which channel coordination takes place (Frazier and Rody 1991). As such, influence strategies are ubiquitous and readily identifiable by supply chain members (Frazier and Summers 1984). When actions are observed, imitation may occur, even when organizations may not intend for this to happen (Greve 1998; Henisz and Delios 2001). Consequently, influence strategies provide a readily identifiable and rich source of behaviors that can be examined in an institutional context to determine the extent to which contagion occurs.

Hypotheses

In this section, we develop hypotheses that predict conditions under which supply chain contagion is more or less likely to occur on the basis of reflexive imitation, compliant imitation, and normative imitation motivations. We describe each of these in turn.

Reflexive Imitation

Environmental uncertainty makes it difficult for firms to select appropriate actions because the potential risks and rewards of any actions are amplified by the uncertainty (Greve 1998; March and Simon 1958; North 1990) and "previous patterns of behavior are less informative" (Johnson, Sohi, and Grewal 2004, p. 24). In situations in which a firm's best course of action is unclear, imitating the behavior of other firms within their network provides a low-cost solution to that uncertainty (Henisz and Delios 2001). Firms are more likely to imitate other firms with which they have direct network ties and interpersonal contact through boundary-spanning personnel (Galaskiewicz and Wasserman 1989). In general, marketing tactics can be an important ingredient for obtaining higher customer-based performance (Zahay and Griffin 2004). Thus, imitating others' tactics may be considered a safe strategy in the presence of environmental uncertainty. Therefore, when environmental uncertainty is present, dealers are more likely to imitate a manufacturer's use of influence strategies with other parties.

Reflexive imitation can also occur under conditions of similarity. Perceived similarity among boundary-spanning personnel increases the likelihood that organizational practices are diffused among those individuals (Galaskiewicz and Wasserman 1989; Strang and Meyer 1994). Individuals who are similar are more likely to identify with one another and to form shared conceptions of social reality (Friedland and Alford 1991). Identity has increasingly been acknowledged as a salient attribute of organizations and organizational forms (Romanelli and Khessina 2005). Through interpretations of situations and events, individuals affect the behavior of their organizations (Lant and Baum 1995; Pfeffer and Salancik 1978). As a result, greater imitation is likely to occur with greater similarity between the firms' individual boundary spanners (Suchman 1995).

H_1: Environmental uncertainty is positively associated with supply chain contagion in the form of dealers' imitation of the manufacturer's use of downstream influence strategies.

H_2: Similarity between manufacturer and dealer boundary-spanning personnel is positively associated with supply chain contagion in the form of dealers' imitation of the manufacturer's use of downstream influence strategies.

Compliant Imitation

DiMaggio and Powell (1983) suggest that interorganizational dependence can increase the likelihood of imitation between organizations. Their primary view, as well as those of other researchers (e.g., Pfeffer and Salancik 1978), is that organizations imitate those they are dependent on to increase their perceived validity by those organizations. We suggest that in the case of interorganizational relationships, cohesiveness between the organizations is paramount in determining the degree of compliant imitation (Galaskiewicz and Burt 1991). The marketing literature suggests that dependence is a bilateral construct that should be studied in terms of interdependence and dependence asymmetry (Heide and John 1988; Kumar, Scheer, and Steenkamp 1998).[2] Interdependence (sometimes referred to as "dependence magnitude"; see Gundlach and Cadotte 1994) is the sum of both organizations' dependence on each other, and dependence asymmetry refers to the difference between the organization's dependence on its partner and the partner's dependence on the organization (Geyskens and Steenkamp 1995).

---

[1] We use more straightforward terminology for what sociologists have referred to as "isomorphism." Specifically, we refer to "mimetic isomorphic" as "reflexive imitation," to "coercive isomorphic" as "compliant imitation," and to "normative isomorphic" as "normative imitation." We thank a knowledgeable reviewer for suggesting this terminology.

The more dependent two organizations are on each other (i.e., greater interdependence or magnitude), the more cohesive the relationship is (Emerson 1962). As a relationship becomes more cohesive, the partners make more investments with each other (Anderson and Weitz 1992) and form deeper common interests (Kumar, Scheer, and Steenkamp 1995), which leads to more cooperative interactions (Dwyer, Schurr, and Oh 1987; Gundlach and Cadotte 1994). As organizations increasingly cooperate with each other, they tend to conform to each other in an attempt to maximize the usefulness of that cooperation (Ostrom 2000). Consequently, the greater the interdependence, the more likely imitation is to occur.

In contrast to the cohesion associated with greater interdependence, higher levels of dependence asymmetry may cause relationships to become unstable and dysfunctional (Geyskens and Steenkamp 1995; Narayandas and Rangan 2004). According to Emerson (1962), the instability surfaces because the less dependent organization may exercise its power advantage in the relationship. Even without evidence of the exercise of power, in an unbalanced relationship, the more dependent organization may attempt to balance the relationship by becoming less dependent (Mindlin and Aldrich 1975). For example, the more dependent organization may seek out relationships with alternative organizations or attempt autonomous acts to reduce its dependency on the less dependent organization (Ulrich and Barney 1984). With increased potential for the disintegration of the relationship, the more dependent organization feels less need to be compliant; cohesiveness is lower, and imitation will be less likely to occur.

H_3: Interdependence is positively associated with supply chain contagion in the form of dealers' imitation of the manufacturer's use of downstream influence strategies.

H_4: Dependence asymmetry is negatively associated with supply chain contagion in the form of dealers' imitation of the manufacturer's use of downstream influence strategies.

Normative Imitation

Because normative prescriptions of behavior develop between firms over time, imitation of behavior between organizations in a given environment increases as the level of their interactions and the interactions of their boundary-spanning personnel increases (DiMaggio and Powell 1983; Scott 2001; Strang and Meyer 1994). Firms and their personnel are motivated to achieve not only economic success but also social legitimacy (Carroll and Delacroix 1982). Specifically, collective beliefs and a common acceptance of legitimacy standards tend to develop as interaction levels increase (Aldrich and Fiol 1994; Leblebici et al. 1991). Such standards are used to make judgments about "whether [a behavior] is 'the right thing to do'" (Suchman 1995, p. 579). As a result, individuals who interact in the same organizational environment come to accept a narrower range of behaviors as the norm to attain or maintain legitimacy. As this range of legitimate behaviors narrows, firms whose personnel interact more frequently are increasingly likely to display matching behaviors (Romanelli and Khessina 2005).

In addition to legitimacy motivations that can arise from these straightforward interactions, the length of tenure of individuals in the same organizational environment will also prescribe behaviors for firms (Larson 1977). As industrial tenure increases, the promulgation of normative rules about organizational behavior that emerges from formal education and other organizational networks (e.g., professional and trade associations) increases. Key boundary-spanning personnel across a range of organizations that operate in the same environment often interact. These key employees tend to have more similar professional beliefs and are predisposed to accept the same organizational behaviors (Perrow 1974). Thus, the longer an employee's tenure in an industrial field, the greater is the likelihood that the employee will follow prescriptive norms. In turn, these key employees influence the norms established in their own firms regarding organizational behavior (Levitt and Nass 1989). As the range of possible behaviors narrows, supply chain contagion is likely to occur to a greater extent.

H_5: Frequency of contact between manufacturer personnel and dealer personnel is positively associated with supply chain contagion in the form of dealers' imitation of the manufacturer's use of downstream influence strategies.

H_6: Industry tenure is positively associated with supply chain contagion in the form of dealers' imitation of the manufacturer's use of downstream influence strategies.

Sample Design

The unit of analysis in this study is the vertical supply chain consisting of three members—that is, manufacturer–dealer–customer triads (see Figure 1). We gained the cooperation of a *Fortune* 500 agricultural equipment manufacturer that sells its products through a dealer network. In this setting, the dealer principal is the primary contact who interacts with the manufacturer's territory manager (TM; each dealership deals with only one TM) and, at the same time, is characteristically the primary person who interacts with the customer (we only include data in the study in which this was the case). Business customers are typically represented by an individual buyer who is responsible for making all purchase decisions. Therefore, the key boundary spanner at each level of the supply chain could be identified and matched on a one-to-one-to-one individual basis.

We obtained contact information from the manufacturer for 400 dealers that were randomly selected within each of the manufacturer's 100 territories (4 per territory) as well as the names and addresses of approximately three customers for each dealer, who were chosen at random from each dealer's customer database (we had accurate contact information for 1188 customers). Each customer made at least one purchase over the previous year. In addition, the manufacturer reported on the territory sales accounted for by each dealership so that interdependence and dependence asymmetry could later be calculated. Dealer principals and customers were simultaneously mailed packets that contained a questionnaire, a postage-paid return envelope, and a signed cover letter that addressed the recipient by name. Figure 2 summarizes the data collection process.

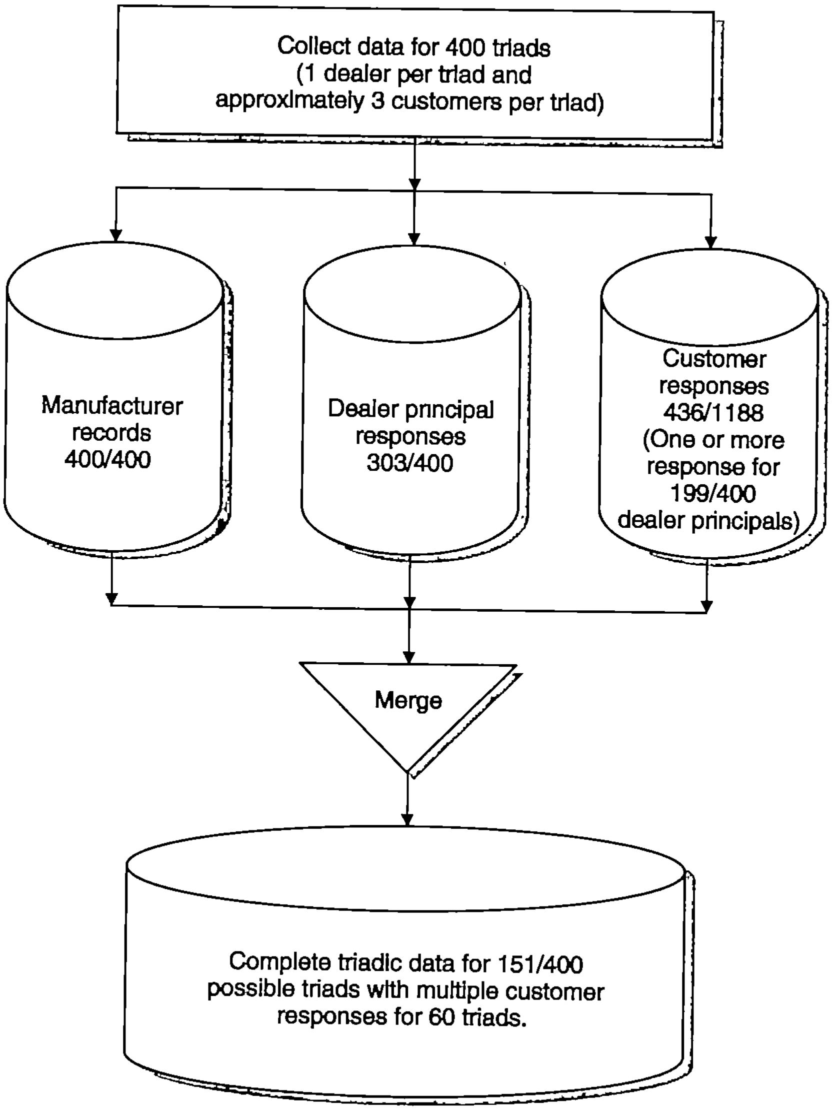
---

[2] We modified and improved this section with the helpful guidance of a knowledgeable reviewer.

After three waves of mailings, we received 307 dealer principal responses; we eliminated 4 because of missing data, which left 303 usable responses. We obtained 436 usable customer responses with complete data. In total, we received one or more customer response that named 199 of the 400 dealer principals. Matching of the manufacturer, dealer principal, and customer responses resulted in 151 unique triads. Of these triads, we had two or more customer responses within 60 triads. As we discuss subsequently, when two or more responses were received for a single dealer principal, we took the average of these multiple responses within each triad. We tested nonresponse bias using a median split between early and late responders for each sample. There were no significant differences between early and late responders in any sample.

Our methodology and sample design are consistent with the recommendations of Podsakoff and colleagues (2003) for reducing common methods biases. They recommend (1) obtaining measures of the predictor and criterion variables from different sources; (2) offering anonymity/confidentiality to reduce the chances of responses that are socially desirable, lenient, acquiescent, and consistent with how

FIGURE 2
Triadic Data Collection Procedure


Collect data for 400 triads
(1 dealer per triad and
approximately 3 customers per triad)
Manufacturer records 400/400
Dealer principal responses 303/400
Customer responses 436/1188 (One or more response for 199/400 dealer principals)
Merge
Complete triadic data for 151/400 possible triads with multiple customer responses for 60 triads.

respondents believe researchers want them to respond; and (3) informing the respondents that there are no correct or incorrect answers and that they should respond as honestly as possible to reduce evaluation apprehension. We followed these recommendations. In addition, our key variables are not self-reported. There are many benefits of non-self-reported measures, including the reduction of self-serving biases (McFarland, Challagalla, and Shervani 2006).

Measures

Table 2 lists the summary statistics and the correlation matrix for all variables. The items, item reliabilities, and factor reliabilities for all multi-item measures appear in the Appendix.

Dependent Variable

We examine supply chain contagion in the form of downstream influence strategy similarity between manufacturer and dealer (A → B) and between dealer and customer (B → C). We model manifest influence strategy contagion (or, more simply, manifest contagion) as a second-order composite latent variable composed of the first-order influence strategy similarity measures (Jarvis, Mackenzie, and Podsakoff 2003) (see Equation 1). We used the influence strategy measures that McFarland, Challagalla, and Shervani (2006) developed. These six measures are applicable to both channel coordination and sales settings. "Information exchange" is a communication process in which information about general business issues is exchanged but no specific requests are made. "Recommendations" are logical arguments that explain the benefits the target will achieve if the target complies with the request of the source of influence. With "promises," the source of influence offers a specific reward or inducement that is contingent on the target's compliance with the source's request. "Threats" are implied or stated negative sanctions or consequences that will occur if the target does not comply with the source's request. "Ingratiation" involves the use of behaviors that are designed to enhance interpersonal attractiveness to gain the approbation of others and includes actions such as praising, acting humble, and so forth. "Inspirational appeals" attempt to elicit an emotional response, such as enthusiasm, and strive to excite the target about the source's request.

We scored all influence strategy items on a five-point scale anchored by never uses and always uses. To eliminate the problems associated with self-reported measures and to reduce common methods biases, we asked the dealers to rate the frequency with which the manufacturer used influence strategies with them (labeled ISM in Equation 1), and we asked the customers to rate the frequency with which the dealers used influence strategies with them (labeled ISD in Equation 1). We determined the similarity of use for each of the six influence strategies as illustrated in Equation 1. Higher scores indicate greater imitative use of influence strategies. Peter, Churchill, and Brown (1993) note that similarity scores should be used with caution because of problems such as the possibility of inflated reliability scores. However, for many research scenarios, their use is appropriate and common in areas such as satisfaction research (e.g., Bohlman et al. 2006), measuring psychic distance and cultural distance in cross-cultural settings (e.g., Dow and Karunaratna 2006), and measuring an individual's propensity to use coercive influence (e.g., Su, Fern, and Ye 2003). Before calculating these scores, we ensured that the influence strategy measures were invariant across groups; we followed the procedure that Steenkamp and Baumgartner (1998) recommend (we discuss this procedure in greater detail subsequently).

$$(1) \qquad y_i = 5 - \left| \mathrm{ISM}_{ij} - \left[\left(\sum_{k=1}^{n} \mathrm{ISD}_{ijk} \right) / n \right] \right|,$$

where

i = influence strategy identifier, 1–6;
j = triad identifier, 1–151; and
n = number of consumer responses for each triad, 1–3.

In the majority of cases, we had single informants for each of the three members of the triad; however, as we discussed previously, for 60 of the 151 triads, we had more than one customer response.[3] The customers report on only the dealers' behaviors; thus, they can be viewed as key informants. As Equation 1 shows, in these cases, we followed convention and aggregated these responses using the mean (e.g., Bell, Menguc, and Stefani 2004; Homburg and Furst 2005; Kumar, Scheer, and Steenkamp 1995, 1998).[4] To ensure that aggregation is appropriate, the within-group variation should be smaller than the between-group variation; in other words, we expect that the data will exhibit a high degree of interrater reliability (James 1982). To explore this issue, we conducted several analytic tests (Bartko 1976; Bliese 2000; James 1982). As Table 3 depicts, the analyses of variance are significant for each of the variables the customers reported, and each eta-square (η^2) exceeds the threshold of .20, indicating that within-group variation is smaller than the between-group variation (Bell, Menguc, and Stefani 2004; Bliese 2000). The interclass correlations (ICC) also support aggregation. The ICC(1) coefficient can be viewed as the consistency of the responses and, in general, should exceed .20. The ICC(2) can be viewed as the reliability of the within-group means (Bell, Menguc, and Stefani 2004; Bliese 2000). Together, this indicates strong support for the legitimacy of aggregating these responses.[5]

Manifest contagion is modeled as a composite latent variable specified as what Jarvis, Mackenzie, and Podsakoff (2003) describe as a Type II, reflective first-order, formative second-order model. The six first-order factors are mea-

---

[3] We thank a knowledgeable reviewer for suggesting this.

[4] As one reviewer noted, alternatives to data aggregation might include repeated measures methodologies. However, the reviewer noted that these techniques are unlikely to yield substantially different results than were achieved by data aggregation in our case.

[5] We conducted these tests with the helpful guidance of a knowledgeable reviewer.

TABLE 2
Descriptive Statistics, Correlation Matrix, and Manifest Contagion Factor Loadings

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Independent Variables														
1. Environmental uncertainty														
2. Similarity	.06													
3. Interdependence	.09	–.05												
4. Dependence asymmetry	–.02	–.04	–.06											
5. Frequency of contact	–.02	.03	–.01	.01										
6. Industry tenure	.13	.13	–.01	.08	.01									
Control Variables														
7. Commitment	–.14	.21	.08	.07	.04	.20								
8. Continuance	–.26	.11	.14	.11	–.01	.07	.71							
Influence Strategy Similarity Variables														
9. Information exchange	–.01	.29	–.01	–.21	.31	–.07	–.10	–.06						
10. Recommendations	–.08	.15	.03	–.16	.01	–.05	–.06	–.07	.45					
11. Promises	.07	.04	–.05	.09	.00	.10	–.13	–.01	.00	–.03				
12. Threats	–.05	.22	.04	–.18	.18	–.20	.04	.02	.22	.06	.27			
13. Ingratiation	–.02	.21	–.03	–.10	.18	.07	–.03	–.05	.39	.31	–.14	–.12		
14. Inspirational appeals	.10	.22	–.17	–.16	.11	–.03	–.09	–.04	.48	.46	–.05	.02	.43	
Manifest contagion^a	.20	.29	.02	–.25	.30	–.09	–.16	–.04	*.70*	*.66*	*.33*	*.55*	*.60*	*.50*
M	3.86	4.10	.95	.34	3.89	19.44	3.60	4.12	4.04	3.78	4.32	4.68	4.03	3.94
SD	.80	.76	.38	.25	.71	11.38	.99	.80	.87	.91	.85	.69	.82	.89

[a]Influence strategy factor loadings onto manifest contagion are in italicized, underlined font based on the partial least squares confirmatory factor analysis.

Notes: $p \leq .05$ for all $r > \pm.16$, $p \leq .01$ for all $r > \pm.21$, and $p \leq .001$ for all $r > \pm.25$.

TABLE 3
Tests for the Appropriateness of Aggregating Multi-Informant Customer Data

Dealers Influence Strategy Usage Reported by Customer	$F_{(69, 70)}$ for Analyses of Variance	η^2	ICC(1)	ICC(2)
Information exchange	3.870**	.77	.59	.74
Recommendations	2.390**	.67	.41	.58
Promises	1.664*	.58	.25	.40
Threats	2.931**	.71	.49	.66
Ingratiation	1.728*	.59	.27	.42
Inspirational appeals	2.688**	.69	.46	.63

*$p < .05$.
**$p < .01$.

sured as in Equation 1. The manifest contagion construct meets the standard definition of a formative construct because it represents a combination of factors that are additive rather than reflective or caused by the latent factor (Fornell and Bookstein 1982). Thus, the model does not assume that similarity in the use of any one influence strategy is caused by a single underlying construct but rather that manifest influence strategy contagion is a linear combination of each influence strategy similarity factor (Diamantopoulos and Winklhofer 2001).

Independent and Control Variables

Environmental uncertainty. This multi-item scale measures the difficulty that firms have in making predictions about future business conditions in their industry. We adapted our measures from Celly and Frazier's (1996) study, using a five-point semantic differential scale. This measure was based on the dealers' responses.

Manufacturer–dealer similarity. This variable measures the perceived similarity of the dealer with the TM from the dealer's perception. We adapted this multi-item scale from the work of Doney and Cannon (1997) and modified it slightly for context; we used a five-point Likert-type scale with scale anchors "strongly disagree" and "strongly agree."

Interdependence and dependence asymmetry. We asked the manufacturer to report the percentage of territory sales accounted for by each dealership in the sample. Similarly, we asked each dealer principal to report the percentage of total purchases made from the focal manufacturer. Following precedent (e.g., Gundlach and Cadotte 1994; Jap and Ganesan 2000; Lusch and Brown 1996) and in line with the condition that dependence should be judged from both sides of a dyad, we determined total dependence by summing these two values, and we determined asymmetry by taking the absolute difference between them (Geyskens and Steenkamp 1995; Kumar, Scheer, and Steenkamp 1995).

Frequency of contact. This variable measures the perceived frequency of contact between the dealer and the TM from the dealer's perspective. We adapted this multi-item measure from the work of Doney and Cannon (1997); we used a five-point Likert-type scale anchored by "strongly disagree" and "strongly agree."

Industry tenure. This is a single-item measure that asks the dealer principals to indicate in years and months the length of time they have been working in their current profession.

Control variables. We measured commitment and expectations of continuing the relationship from the dealers' perspective with regard to the manufacturer. We included these relationship variables because of their potential impact on behaviors and perceptions in the supply chain. We based these multi-item scales on the work of Kumar, Scheer, and Steenkamp (1995). These measures use five-point Likert-type scales with scale anchors of "strongly disagree" and "strongly agree."

Measurement Analysis

We analyzed one confirmatory factor analysis (CFA) model for all multi-item measures reported by dealers and one CFA for the customer respondents using maximum likelihood estimation in LISREL 8.54 (Jöreskog and Sörbom 1996). The models demonstrated good fit. For the first model, $\chi^2 = 748$, with 610 degrees of freedom; root mean square error of approximation (RMSEA) = .039; nonnormed fit index (NNFI) = .96; comparative fit index (CFI) = .96; and incremental fit index (IFI) = .96. For the second model, $\chi^2 = 367$, with 184 degrees of freedom; RMSEA = .067; NNFI = .95; CFI = .96; and IFI = .96. We also conducted a CFA on the second-order formative measure, manifest contagion; we discuss these results subsequently.

For all multi-item measures, composite trait reliabilities exceeded the recommended cutoff criterion of .70, and the variance extracted for each factor exceeded the .50 cutoff criterion that Fornell and Larcker (1981) suggest (except for ingratiation, which had a variance extracted of .49), indicating that the measures are internally consistent. All item loadings were significant for each factor, demonstrating convergent validity (Anderson and Gerbing 1988). We demonstrated discriminant validity following the procedures of Fornell and Larcker (1981), and we demonstrated unidimensionality following the guidelines of Gerbing and Anderson (1988).

Grewal, Cote, and Baumgartner (2004) highlight the problems that the presence of multicollinearity can create when structural equation modeling procedures are used. They recommend checking for multicollinearity by examining the strength of correlations between constructs and conducting Fornell and Larcker's (1981) test for discriminant validity. It does not appear that multicollinearity is a threat to the validity of the current results.

Test of Measurement Invariance Across Response Groups

When difference scores between groups are calculated, the results are meaningful only if the measures are invariant across groups (Steenkamp and Baumgartner 1998). Thus, we tested for measurement invariance for the influence strategies across groups. Steenkamp and Baumgartner recommend testing for both configural invariance and metric invariance when similarity scores are calculated. Configural invariance exists when the measures have the same factor structure across groups. Testing this involves an examination of the degree to which the observed variables fit the latent constructs in each group, allowing factor loadings to vary freely within each group. Metric invariance exists when both the factor structure and the scale of each observed variable are equivalent. Testing metric invariance involves an examination of the equality of a construct's metrics across groups through the use of hierarchical nested multigroup CFA models, in which the fit statistics of the baseline unconstrained configural invariance model are examined and then compared with the fit statistics of the constrained metric invariance model. For the baseline configural invariance model, $\chi^2 = 703$, with 388 degrees of freedom; CFI = .9562; and gamma hat = .9080, which indicates that the model has acceptable fit and meets the test for configural invariance. For the constrained metric invariance model, $\chi^2 = 728$, with 404 degrees of freedom; CFI = .9510; and gamma hat = .9067. The differences in the fit indexes for the hierarchical nested models are $\Delta\chi^2 = 25$, with 16 degrees of freedom (nonsignificant at $p = .05$); ΔCFI = –.0052; and Δgamma hat = –.0069. This indicates that the null hypothesis of metric invariance should not be rejected (Steenkamp and Baumgartner 1998). Thus, the influence strategy measures demonstrate both configural invariance and metric invariance across the dealer and the customer samples, and difference scores are meaningful.

Results

Test of Hypotheses

We tested the hypotheses using partial least squares (PLS) structural equation modeling with PLS-Graph 3.0 (Chin 2001). Partial least squares is a non-covariance-based structural equation modeling procedure (Fornell and Bookstein 1982) commonly used in the marketing literature (e.g., Hennig-Thurau et al. 2006; Johnson, Herrmann, and Huber 2006; Smith and Barclay 1997). A strength of PLS is its ability to incorporate formative scales, as does our model (see Dellande, Gilly, and Grahm 2004; Ulaga and Eggert 2006; White, Varadarajan, and Dacin 2003). Because PLS is a nonparametric procedure (Wold 1985), the structural model is assessed using the R-square for dependent constructs and the size, t-statistics, and significance level of the structural path coefficients. Standard errors and t-statistics are estimated using the bootstrap resampling procedure with a resample size of 200, as Chin (1998) recommends.

Initially, we conducted a CFA for the higher-order formative construct manifest contagion, using the six influence strategy similarities factors as first-order constructs with PLS. We used the hierarchical component procedure that Wold (1985) recommends for estimating the second-order factor (Chin 1998). The composite reliability is .73. All first-order factors load significantly on the higher-order construct ($p < .001$), indicating convergent validity (see Table 2). We indicate discriminant validity following the procedures of Fornell and Larcker (1981).

Next, we estimated the structural model. Table 4 lists the path loadings and their t-statistics. Because the promises influence strategy did not load significantly on manifest contagion in the full structural model, we dropped this factor from the model. For each multi-item factor, the composite reliability is ≥.86, and the average variance extracted is ≥.61. The R-square for the model is .30. Four of the six hypotheses are supported. H_1 predicted a positive relationship between environmental uncertainty and manifest contagion; the standardized path coefficient (SPC) was .202

TABLE 4
Results of PLS Analysis on the Dependent Variable Manifest Contagion

	Hypothesis	Imitation Motivation	Predicted Direction	Standardized Path Loading	Hypothesis Supported?
Independent Variable					
Environmental uncertainty	H_1	Reflexive	+	.202*	Yes
Manufacturer–dealer similarity	H_2	Reflexive	+	.286**	Yes
Interdependence	H_3	Compliant	+	–.001	No
Dependence asymmetry	H_4	Compliant	–	–.233*	Yes
Frequency of contact	H_5	Normative	+	.270**	Yes
Industry tenure	H_6	Normative	+	–.097	No
Control Variable					
Commitment	Control			–.021	N.A.
Continuance	Control			–.061	N.A.

*$p < .05$.
**$p < .01$.
Notes: N.A. = not applicable.

($p < .05$), indicating support for the hypothesis. H_2 predicted a positive relationship between manufacturer–dealer similarity and manifest contagion; this hypothesis is also supported (SPC = .286, $p < .01$). Interdependence (H_3) was not significantly related to manifest contagion; however, dependence asymmetry (H_4) was negatively related to manifest contagion, as we hypothesized (SPC = –.233, $p < .05$). H_5 predicted a positive relationship between the frequency of contact between manufacturer–dealer and manifest contagion; this hypothesis is also supported (SPC = .270, $p < .01$). Industry tenure (H_6) was not significantly related to manifest contagion.

Summary of Findings

The results of the conclusive study support the conclusions we drew from our in-depth field interviews. That is, although the literature has not discussed the idea of supply chain contagion, we show that intermediaries often imitate the behaviors of their suppliers. One interviewee even referred to this as the "matching principal." A large variance ($R^2 = .30$) in the use of downstream influence strategies is explained as supply chain contagion. An implicit assumption that firms are likely to make is that the effects of influence strategy behaviors are confined to the dyadic relationship. Given the extensive and sustained attention in the literature on marketing influence strategies (e.g., Boyle and Dwyer 1995; Boyle et al. 1992; Frazier and Rody 1991; Frazier and Summers 1984; McFarland, Challagalla, and Shervani 2006; Payan and McFarland 2005), the explanatory power of supply chain contagion is impressive and suggests that this newly identified phenomenon is important. This phenomenon may also occur with other firm behaviors in addition to influence strategies.

In addition to supporting the existence of supply chain contagion, there is solid support for our conceptual framework. The two reflexive imitation hypotheses were supported, and one of two hypotheses were supported for both compliant and normative imitation pressures. To reiterate, H_1 (environmental uncertainty) and H_2 (similarity) were tests of reflexive imitation, H_3 (interdependence) and H_4 (dependence asymmetry) were tests of compliant imitation, and H_5 (frequency of contact) and H_6 (industry tenure) were tests of normative imitation. Each of these factors might influence contagion. However, the influence of these factors may vary depending on situational conditions. As we predicted, in our sample, environmental uncertainty and similarity characteristics (reflexive imitation) as well as frequency of contact (normative imitation) had a positive influence on dealers imitating manufacturers, whereas dependence asymmetry (compliant imitation) had a negative influence.

Discussion

Management of the supply chain continues to receive increasing attention as organizations look for ways to improve the creation, production, distribution, and marketing of products (e.g., Lambert and Cooper 2000; Lee, So, and Tang 2000; Schroeder, Bates, and Junttila 2002). Supply chains may contain many organizational levels, but research has predominantly focused on individual dyads when developing theory for supply chain relationships (Wathne and Heide 2004). Although a dyadic focus is important in the process of theory development, the limitation of taking a purely dyadic focus is that critical factors within the system or network in which organizations operate are not in a researcher's consideration set. A strength of this study is the use of data from all three levels of the supply chain. Our research approach enables us to identify a unique phenomenon that would not be possible without the triadic study design. Focusing on the supply chain, instead of the status quo of dyadic, reciprocal interactions between firms, enables us to identify and support empirically the existence of supply chain contagion (i.e., the spread of interfirm behaviors from one dyad to an adjacent dyad in the supply chain). Our model suggests that how intermediaries treat end customers is explained to a great extent as the intermediaries simply imitating how they were treated themselves by their suppliers.

Our theoretical framework, which predicts the antecedents that make supply chain contagion more or less likely to occur, has an institutional theoretical foundation. Institutional theory explicitly acknowledges that the effects of firm behavior are not confined to dyadic relationships and that the behaviors of organizations are influenced by their larger social context. Recognizing the importance of the institutional environment as a determinant of channel behaviors, processes, and structures, Grewel and Dharwadkar (2002) call for empirical research on key aspects of the institutional environment in channels of distribution. Our research is consistent with and builds on Grewel and Dharwadkar's theoretical article. Institutional theory has been increasingly acknowledged as important in explaining firm behavior (e.g., Grewal, Comer, and Mehta 2001; Handelman and Arnold 1999; Homburg, Workman, and Krohmer 1999; Srinivasan, Lilien, and Rangaswamy 2002), and it has particular relevance in the context of this study; it suggests that organizations operate within a social network and that imitation between organizations occurs under (1) reflexive conditions of environmental uncertainty and similarity and (2) normative conditions of contact frequency but is less likely to occur under (3) compliant (or noncompliant) conditions of dependence asymmetry.

This study makes several contributions. First, this study develops the supply chain contagion concept, which, to the best of our knowledge, is a novel concept. As a result, this study contributes to the understanding of how firm behaviors become imitated in the supply chain. Second, through the use of institutional theory, we described these effects as stemming from direct interactions between firms. Notably, these interactions can stimulate the participants to imitate one dyadic partner when dealing with other partners in future interactions. Our field interviews suggest that this imitation is typically not part of the planned response by dyadic partners; nonetheless, it may have both positive and negative implications. Thus, contagion may occur inadvertently, and it may occur both with or without the knowledge of the involved party. Our field interviews also suggest that supply chain contagion can occur with wholesalers, retailers, and service providers. Third, we highlight the importance of institutional theory in providing a more complete understanding of the behavior of firms in the supply chain. From our sample, we found that supply chain members' behavior was influenced by institutional pressures. Specifically, in some cases, the manufacturer's interactions with dealers induced dealers to imitate its behaviors. Institutional theory provides a convincing explanation for the social environmental effects on firm behavior. Because firm behavior occurs in a social environment, institutional theory can advance the understanding of firm behavior in the supply chain beyond simple dyadic-based explanations. Fourth, we focus on the importance of considering both macroenvironmental and microenvironmental institutions in explaining behaviors within the supply chain. The results of our empirical study lend support to this approach. Finally, we used influence strategies as a proxy for determining the level of contagion within a supply chain, so our results have important implications for influence strategy researchers. Until now, influence strategy research has focused largely on bilateral or unilateral antecedents of influence strategies. Our model suggests that effects beyond the dyad are critical determinants of the use of influence strategies within channels of distribution. In summary, these contributions demonstrate the importance of looking outside the focal dyad when examining certain phenomena.

Limitations and Further Implications for Research

Several limitations of this study may affect the generalizability of its results. The characteristics of the market in which our study participants operate may not be equivalent in other industries. The particular mix of influence strategies used likely varies across industries; therefore, our findings may not generalize to other industries. Continued research in a variety of industries might help determine industry-specific effects. Although influence strategies are perhaps the strongest and most logical choice for examining supply chain contagion, other dependent variables might prove useful in the assessment of these effects in the supply chain.

We used a general measure of environmental uncertainty in this study (e.g., Celly and Frazier 1996). However, although we found support for the effect of environmental uncertainty on contagion, environmental uncertainty has also been identified as a multidimensional construct in the literature. Some examples include demand or volume uncertainty, technological uncertainty, goal ambiguity, and performance ambiguity (see, e.g., Grewal and Tansuhaj 2001; Heidi and John 1990). To provide additional clarity, further research should examine the effects of these different forms of uncertainty on contagion. In addition, a more direct measure of professionalism than industry tenure might be more appropriate. Although we found frequency of contact to be a significant normative imitation pressure, its companion measure, industry tenure, was not. DiMaggio and Powell (1983) discuss professionalism in the context of membership in associations, participation in industry-sponsored training programs, and so forth. Future studies should examine whether alternative variables for professionalism are antecedents to contagion. Some scholars have also pointed to problems associated with similarity/difference scores; thus, examining the phenomena of supply chain contagion without the use of difference scores is important to provide additional empirical support for our findings.

We reiterate that we chose the six factors investigated in this study as predictors of supply chain contagion on the basis of institutional theory. We must also acknowledge that other, noneconomic factors might influence the extent of contagion within the supply chain. For example, a potential environmental characteristic could be the percentage of other firms that already engage in imitated behaviors (Greve 1996; Henisz and Delios 2001). Firm characteristics, such as age, size, and strategy, may influence relationships, as institutional theory posits (Guillén 2002). Older firms might be able to rely more on experience and might be more prone to habitual behavior than newer firms. Larger firms might be less likely to imitate smaller firms. In addition, there may be an array of interactions among these factors that could shed light on supply chain contagion.

Outcomes of supply chain contagion are also of interest. Supply chain contagion may occur for various reasons and without the recognition of organizations or the resultant performance implications. To assess the nature of the relationship between supply chain contagion and organizational performance, researchers could examine multiple industries with a variety of different characteristics. In addition, attention could be given to the specific types of organizations being imitated. For example, competitors, suppliers, and government organizations could all be sources of behavior subject to imitation. Just how far contagion carries is also of importance to researchers. Potentially, both multilevel contagion and contagion in lateral and network relationships may exist. Researchers may find a reduction in supply chain contagion at each additional level of the supply chain in some industries and a stable or increased rate in other industries. In addition, although this study examined supply chain contagion down the supply chain, a similar process may occur moving upstream. A notable implication of the study is the importance of placing dyadic encounters within a larger context. Dyadic interactions are still worthy of study because factors within the dyadic interaction still directly influence the outcomes of the interaction. However, examination of the effects beyond the dyad has shown the existence of more complex phenomena than is frequently assumed.

Managerial Implications

Our findings should be of some assistance to managers as well. In general, managers should attend more to the long-term outcomes of their organizations' actions, which can have ongoing effects by influencing other organizations to act similarly. Whether or not managers desire their organizations' actions to be imitated and passed on, these actions may occur nonetheless. By acknowledging that supply chain contagion might occur, managers should be better able to structure their behaviors to provide the best chance for imitation when it is desired and the least chance for imitation when it is not. For example, by increasing exposure of their organizations' behaviors to other organizations, managers can assist other organizations in copying the behaviors. Alternatively, hiding the behaviors can help prevent them from being imitated. Manufacturers and suppliers may be more hesitant to use coercive influence strategies if they will propagate down the channel because their use may be detrimental to downstream relationships (Payan and McFarland 2005).

Our in-depth field interviews lead us to the conclusion that when contagion occurs, it may lead to suboptimal actions. Institutional theory suggests that contagion can occur without firms' knowledge, and our field interviews appeared to confirm this. There are several notable issues that arise when the behaviors firms engage in within a dyadic interaction are passed on by their dyadic partner. An obvious issue is that firms may need to evaluate the likelihood that they themselves are engaging in inappropriate imitation of channel partner behaviors. Because firms may lose control of the spread of behaviors, they should consider in advance whether it is in their best interests for the behavior to spread in this manner. Managers and boundary-spanning personnel who are aware of supply chain contagion effects may be better inoculated against their own unintended imitation of other organizations within their supply chain. Competitors may also be able to capitalize on an increased awareness and understanding of this process.

Conclusions

This study answers Grewal and Dharwadkar's (2002) call for research that examines firm behavior in the supply chain due to institutional environmental effects. In doing so, this study breaks new ground in several areas and sheds light on some important topics. We found that there are indeed contagion effects from the manufacturer–dealer dyad to the dealer–customer dyad. Manufacturers and dealers should make themselves aware of these effects and more carefully monitor the influence strategies (and other behaviors) they use. For example, manufacturers may believe that it is appropriate to use coercive influence strategies, such as threats, with their dealers, but it is unlikely that manufacturers will desire their dealers to use these same strategies with the final customer (Gundlach and Cadotte 1994). Acknowledging that supply chain contagion occurs in a manner that is consistent with the principles of institutional theory provides important implications for further supply chain research and practice.

APPENDIX
Multi-Item Scales

	Standardized Item Factor Loading		Composite Reliability (Variance Extracted)	
	Dealer	Customer	Dealer	Customer
Influence Strategies Scale Items[a] (Manufacturers' TMs' Use of Influence Strategies with Dealers Reported by Dealers/Dealers' Use of Influence Strategies with Customers Reported by Customers)				
Information Exchange			.86 (.62)	.84 (.57)
The TM/dealer ...				
1. Asked about your long-term purchasing goals.	.64	.71		
2. Presented information related to your various purchase options.	.87	.67		
3. Asked if there were any problems or needs he or she could help address.	.79	.71		
4. Talked about the possible applications of his or her products or services.	.82	.90		
Recommendations			.89 (.67)	.89 (.67)
The TM/dealer ...				
1. Made it clear that by following his or her recommendations, your business would benefit.	.75	.85		
2. Outlined the evidence for expecting success from his or her suggestions.	.73	.88		
3. Provided a clear picture of the positive impact on your business a recommended course of action would have.	.88	.80		
4. Made a logical argument supporting his or her specific suggestions.	.91	.73		
Promises			.88 (.65)	.85 (.59)
The TM/dealer ...				
1. Made promises to give something back for complying with his or her request (e.g., discounts, quicker delivery).	.69	.76		
2. Offered additional benefits for your business after you had been initially reluctant to agree to his or her terms.	.90	.88		
3. Offered to provide incentives to your business for agreeing to his or her purchase requests.	90	.90		
4. Offered a specific deal for your business to change your position on certain issues.	.72	.45		

APPENDIX
Continued

	Standardized Item Factor Loading		Composite Reliability (Variance Extracted)	
	Dealer	Customer	Dealer	Customer
Threats				
The TM/dealer ...			.93 (.76)	.89 (.67)
1. Stated that you would stop receiving preferential treatment if his or her proposals were ignored.	.66	.98		
2. Advised that he or she would stop doing business with you/your company if his or her requests were not followed.	.92	.92		
3. Threatened to become uncooperative if you failed to agree to his or her demands.	.96	.75		
4. Indicated he or she could “make things difficult” for your business if his or her requests were not met.	.92	.56		
Ingratiation				
The TM/dealer ...			.74 (.49)	.75 (.50)
1. Complimented and praised your achievements.	.73	.72		
2. Discussed shared interests and/or hobbies prior to discussing sales issues.	.64	.75		
3. Made you feel good about yourself before making his or her sales pitch.	.72	.64		
Inspirational Appeals				
The TM/dealer ...			.82 (.61)	.76 (.51)
1. Described the use of his or her products or services with enthusiasm and conviction.	.86	.78		
2. Tried to get you excited about what he or she was selling.	.88	.75		
3. Appealed to your values and ideals when asking for your business.	.56	.61		

	Standardized Item Factor Loading	Composite Reliability (Variance Extracted)
Independent and Control Variable Scale Items		
Environmental Uncertainty		.80 (.50)
(Five-Point Semantic Differential Scale. Dealer Responses.)		
How would you describe the *market* environment in which *your* business operates?		
1. “Predictable/unpredictable”	.71	
2. “Easy to monitor trends/difficult to monitor trends”	.58	
3. “Sales volume is stable/sales volume is volatile”	.80	
4. “Sales forecasts are quite accurate/sales forecasts are quite inaccurate”	.71	
Manufacturer–Dealer Similarity[b]		.88 (.70)
(Five-Point Likert-Type Scale. Dealer Responses.)		
The TM ...		
1. Shares similar interests with me.	.86	
2. Is very similar to me.	.84	
3. Has similar values to mine.	.81	
Frequency of Contact with Manufacturer[b]		.86 (.67)
(Five-Point Likert-Type Scale. Dealer Responses.)		
The TM ...		
1. Often visits my place of business.	.71	
2. Makes more calls than other suppliers.	.74	
3. Frequently makes sales calls.	.98	
Commitment to the Manufacturer[b]		.92 (.78)
(Five-Point Likert-Type Scale. Dealer Responses.)		
1. Even if we could, we would not drop this supplier because we like being associated with them.	.82	
2. We want to continue as a customer of this firm, because we genuinely enjoy our relationship with them.	.94	
3. Our positive feelings toward this company are a major reason we continue to work with them.	.89	

APPENDIX
Continued

	Standardized Item Factor Loading	Composite Reliability (Variance Extracted)
Expectations of Continuing the Relationship with the Manufacturer[b] (Five-Point Likert-Type Scale. Dealer Responses.)		.79 (.56)
1. We expect the relationship with this company to continue for a long time.	.85	
2. Renewing the relationship with this company is virtually automatic.	.79	
3. In the next two years, we are likely to terminate this relationship. (reverse coded)	.59	

[a] Items were measured on a five-point Likert-type scale anchored by 1 = "never" and 5 = "always."
[b] Scale anchors are 1 = "strongly disagree" and 5 = "strongly agree."

REFERENCES

Aldrich, Howard E. and C. Marlene Fiol (1994), "Fools Rush In? The Institutional Context of Industry Creation," *Academy of Management Review*, 19 (4), 645–70.

Anderson, Erin and Barton Weitz (1992), "The Use of Pledges to Build and Sustain Commitment in Distribution Channels," *Journal of Marketing Research*, 29 (February), 18–34.

Anderson, James C. and David W. Gerbing (1988), "Structural Equation Modeling in Practice: Review and Recommended Two-Step Approach," *Psychological Bulletin*, 103 (May), 411–23.

———, Håkan Håkansson, and Jan Johanson (1994), "Dyadic Business Relationships Within a Business Network Context," *Journal of Marketing*, 58 (October), 1–15.

Bartko, John J. (1976), "On Various Intraclass Correlation Reliability Coefficients," *Psychological Bulletin*, 83 (5), 762–65.

Bell, Simon J., Bulent Menguc, and Sara L. Stefani (2004), "When Customers Disappoint: A Model of Relational Internal Marketing and Customer Complaints," *Journal of the Academy of Marketing Science*, 32 (1), 112–26.

Bliese, Paul D. (2000), "Within-Group Agreement, Non-Independence, and Reliability: Implications for Data Aggregation and Analysis," in *Multilevel Theory, Research, and Methods in Organizations*, Katherine J. Klein and Steve W. Kozlowski, eds. San Francisco: Jossey-Bass, 349–81.

Bohlman, Jonathan D., José Antonio Rosa, Ruth N. Bolton, and William J. Qualls (2006), "The Effect of Group Interactions on Satisfaction Judgments: Satisfaction Escalation," *Marketing Science*, 25 (4), 301–321.

Boyle, Brett F. and F. Robert Dwyer (1995), "Power, Bureaucracy, Influence, and Performance," *Journal of Business Research*, 32 (3), 189–201.

———, ———, Robert A. Robicheaux, and James T. Simpson (1992), "Influence Strategies in Marketing Channels: Measures and Use in Different Relationship Structures," *Journal of Marketing Research*, 29 (November), 462–73.

Carroll, Glen R. and Jacques Delacroix (1982), "Organizational Mortality in the Newspaper Industries of Argentina and Ireland: An Ecological Approach," *Administrative Science Quarterly*, 27 (2), 169–98.

Celly, Kirti Sawhney and Gary L. Frazier (1996), "Outcome-Based and Behavior-Based Coordination Efforts in Channel Relationships," *Journal of Marketing Research*, 33 (May), 200–210.

Chatterjee, Debabroto, Rajdeep Grewal, and V. Sambamurthy (2002), "Shaping Up for E-Business: Institutional Enablers of the Organizational Assimilation of Web Technologies," *MIS Quarterly*, 26 (June), 65–89.

Chin, Wynne W. (1998), "The Partial Least Squares Approach to Structural Equation Modeling," in *Modern Methods for Business Research*, G.A. Marcoulides, ed. Mahwah, NJ: Lawrence Erlbaum Associates, 295–36.

——— (2001), *PLS-Graph User's Guide: Version 3.0*. Houston: Software Modeling Inc.

Dacin, M. Tina, Christine Oliver, and Jean-Paul Roy (2007), "The Legitimacy of Strategic Alliances: An Institutional Perspective," *Strategic Management Journal*, 28 (2), 169–87.

Dellande, Stephanie, Mary C. Gilly, and John L. Graham (2004), "Gaining Compliance and Losing Weight: The Role of the Service Provider in Health Care Services," *Journal of Marketing*, 68 (July), 78–91.

Diamantopoulos, Adamantios and Heidi M. Winklhofer (2001), "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research*, 38 (May), 269–77.

DiMaggio, Paul J. and Walter W. Powell (1983), "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review*, 48 (April), 147–60.

——— and ——— (1991), "Introduction," in *The New Institutionalism in Organizational Analysis*, Walter W. Powell and Paul J. DiMaggio, eds. Chicago: University of Chicago Press, 1–38.

Doney, Patricia M. and Joseph P. Cannon (1997), "An Examination of the Nature of Trust in Buyer–Seller Relationships," *Journal of Marketing*, 61 (April), 35–51.

Dow, Douglas and Amal Karunaratna (2006), "Developing a Multidimensional Instrument to Measure Psychic Distance Stimuli," *Journal of International Business Studies*, 37 (5), 578–602.

Dwyer, F. Robert, Paul H. Schurr, and Sejo Oh (1987), "Developing Buyer–Seller Relationships," *Journal of Marketing*, 51 (April), 11–27.

Emerson, Richard M. (1962), "Power-Dependence Relations," *American Sociological Review*, 27 (1), 31–41.

Fornell, Claes and Fred L. Bookstein (1982), "A Comparative Analysis of Two Structural Equation Models: LISREL and PLS Applied to Market Data," in *A Second Generation of Multivariate Analysis*, Vol. 1, C. Fornell, ed. New York: Praeger, 289–324.

——— and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 19 (February), 39–50.

Frazier, Gary and Raymond Rody (1991), "The Use of Influence Strategies in Interfirm Relationships in Industrial Product Channels," *Journal of Marketing*, 55 (January), 52–69.

——— and John Summers (1984), "Interfirm Influence Strategies and Their Application Within Distribution Channels," *Journal of Marketing*, 48 (Summer), 43–55.

Friedland, Roger and Robert R. Alford (1991), "Bringing Society Back in: Symbols, Practices, and Institutional Contradictions," in *The New Institutionalism in Organizational Analysis*, Walter W. Powell and Paul J. DiMaggio, eds. Chicago: University of Chicago Press, 232–63.

Frumkin, Peter and Gabriel Kaplan (2006), "Institutional Theory and the Micro-Macro Link," working paper, Kennedy School of Government, Harvard University.

Galaskiewicz, Joseph and Ronald S. Burt (1991), "Interorganization Contagion in Corporate Philanthropy," *Administrative Science Quarterly*, 36 (1), 88–105.

——— and Stanley Wasserman (1989), "Mimetic Processes Within an Interorganizational Field: An Empirical Test," *Administrative Science Quarterly*, 34 (3), 454–79.

Gerbing, David W. and James C. Anderson (1988), "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment," *Journal of Marketing Research*, 25 (May), 186–92.

Geyskens, Inge and Jan-Benedict Steenkamp (1995), "An Investigation into the Joint Effects of Trust and Interdependence on Relationship Commitment," in *Proceedings of the 24th Conference of the European Marketing Academy*, Vol. 24, M. Bergadaa, ed. Paris: European Marketing Academy, 351–71.

Granovetter, Mark (1985), "Economic Action and Social Structure: The Problem of Embeddedness," *American Journal of Sociology*, 91 (3), 481–510.

Greve, Henrich R. (1996), "Patterns of Competition: The Diffusion of a Market Position in Radio Broadcasting," *Administrative Science Quarterly*, 41 (1), 29–60.

——— (1998), "Managerial Cognition and the Mimetic Adoption of Market Positions: What You See Is What You Do," *Strategic Management Journal*, 19 (10), 967–88.

Grewal, Rajdeep, James M. Comer, and Raj Mehta (2001), "An Investigation into the Antecedents of Organizational Participation in Business-to-Business Electronic Markets," *Journal of Marketing*, 65 (July), 17–33.

———, Joseph A. Cote, and Hans Baumgartner (2004), "Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing," *Marketing Science*, 23 (4), 519–29.

——— and Ravi Dharwadkar (2002), "The Role of the Institutional Environment in Marketing Channels," *Journal of Marketing*, 66 (July), 82–97.

———, Gary L. Lilien, and Girish Mallapragada (2006), "Location, Location, Location: How Network Embeddedness Affects Project Success in Open Source Systems," *Management Science*, 52 (7), 1043–1056.

——— and Patriya Tansuhaj (2001), "Building Organizational Capabilities for Managing Economic Crisis: The Role of Market Orientation and Strategic Flexibility," *Journal of Marketing*, 65 (April), 67–80.

Guillén, Mauro F. (2002), "Structural Inertia, Imitation, and Foreign Expansion: South Korean Firms and Business Groups in China, 1987-1995," *Academy of Management Journal*, 45 (3), 509–525.

Gundlach, Gregory T. and Ernest R. Cadotte (1994), "Exchange Interdependence and Interfirm Interaction: Research in a Simulated Channel Setting," *Journal of Marketing Research*, 31 (November), 516–32.

Handelman, Jay M. and Stephen J. Arnold (1999), "The Role of Marketing Actions with a Social Dimension: Appeals to the Institutional Environment," *Journal of Marketing*, 63 (July), 33–48.

Heide, Jan B. and George John (1988), "The Role of Dependence Balancing in Safeguarding Transaction Costs," *Journal of Marketing*, 52 (January), 20–35.

——— and ——— (1990), "Alliance in Industrial Purchasing: The Determinants of Joint Action in Buyer–Supplier Relationships," *Journal of Marketing Research*, 27 (February), 24–36.

——— and ——— (1992), "Do Norms Matter in Marketing Relationships?" *Journal of Marketing*, 56 (April), 32–44.

Henisz, Witold J. and Andrew Delios (2001), "Uncertainty, Imitation, and Plant Location: Japanese Multinational Corporations, 1990-1996," *Administrative Science Quarterly*, 46 (3), 443–75.

Hennig-Thurau, Thorsten, Markus Groth, Michael Paul, and Dwayne D. Gremler (2006), "Are All Smiles Created Equal? How Emotional Contagion and Emotional Labor Affect Service Relationships," *Journal of Marketing*, 70 (July), 58–73.

Homburg, Christian and Andreas Furst (2005), "How Organizational Complaint Handling Drives Customer Loyalty: An Analysis of the Mechanistic and the Organic Approach," *Journal of Marketing*, 69 (July), 95–114.

———, John P. Workman, and Harley Krohmer (1999), "Marketing's Influence Within the Firm," *Journal of Marketing*, 63 (April), 1–17.

Ibarra, Herminia, Martin Kilduff, and Wenpin Tsai (2005), "Zooming In and Out: Connecting Individuals and Collectivities at the Frontiers of Organizational Network Research," *Organizational Science*, 16 (4), 359–71.

James, Lawrence R. (1982), "Aggregation Bias in Estimates of Perceptual Agreement," *Journal of Applied Psychology*, 67 (2), 219–29.

Jap, Sandy D. and Shankar Ganesan (2000), "Control Mechanisms and the Relationship Life Cycle: Implications for Safeguarding Specific Investments and Developing Commitment," *Journal of Marketing Research*, 37 (May), 227–46.

Jarvis, Cheryl Burke, Scott B. Mackenzie, and Philip M. Podsakoff (2003), "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research*, 30 (September), 199–218.

Jepperson, Ronald L. (1991), "Institutions, Institutional Effects, and Institutionalism," in *The New Institutionalism in Organizational Analysis*, Walter W. Powell and Paul J. DiMaggio, eds. Chicago: University of Chicago Press, 143–63.

Johnson, Jean L., Ravi Sohi, and Rajdeep Grewal (2004), "The Role of Relational Knowledge Stores in Interfirm Partnerships," *Journal of Marketing*, 68 (July), 21–36.

Johnson, Michael D., Andreas Herrmann, and Frank Huber (2006), "The Evolution of Loyalty Intentions," *Journal of Marketing*, 70 (April), 122–32.

Jöreskog, Karl G. and Dag Sörbom (1996), *LISREL 8: User's Reference Guide*. Chicago: Scientific Software International Inc.

Kumar, Nirmalya, Lisa Scheer, and Jan-Benedict E.M. Steenkamp (1995), "The Effects of Perceived Interdependence on Dealer Attitudes," *Journal of Marketing Research*, 32 (August), 348–56.

———, ———, and ——— (1998), "Interdependence, Punitive Capability, and the Reciprocation of Punitive Actions in Channel Relationships," *Journal of Marketing Research*, 35 (May), 225–35.

Lambert, Douglas M. and Martha C. Cooper (2000), "Issues in Supply Chain Management," *Industrial Marketing Management*, 29 (1), 65–83.

Lant, Theresa K. and Joel A.C. Baum (1995), "Cognitive Sources of Socially Constructed Competitive Groups: Examples from the Manhattan Hotel Industry," in *The Institutional Construction of Organizations*, W. Richard Scott and Søren Christensen, eds. Thousand Oaks, CA: Sage Publications, 15–38.

Larson, Magali Sarfatti (1977), *The Rise of Professionalism: A Sociological Analysis*. Berkeley: University of California Press.

Leblebici, Huseyin, Gerald R. Salancik, Anne Copay, and Tom King (1991), "Institutional Change and the Transformation of Interorganizational Fields: An Organizational History of the U.S. Radio Broadcasting Industry," *Administrative Science Quarterly*, 36 (3), 333–63.

Lee, Hau L., Kut C. So, and Christopher S. Tang (2000), "The Value of Information Sharing in a Two-Level Supply Chain," *Management Science*, 46 (5), 626–43.

Levitt, Barbara and Clifford Nass (1989), "The Lid on the Garbage Can: Institutional Constraints on Decision Making in the Technical Core of College-Text Publishers," *Administrative Science Quarterly*, 34 (2), 190–207.

Lusch, Robert F. and James R. Brown (1996), "Interdependency, Contracting, and Relational Behavior in Marketing Channels," *Journal of Marketing*, 60 (October), 19–38.

March, James G. and Johan P. Olsen (1984), "The New Institutionalism: Organizational Factors in Political Life," *The American Political Science Review*, 78 (September), 734–49.

——— and Herbert A. Simon (1958), *Organizations*. New York: John Wiley & Sons.

McFarland, Richard G., Goutam N. Challagalla, and Tasadduq A. Shervani (2006), "Influence Tactics for Effective Adaptive Selling," *Journal of Marketing*, 70 (October), 103–117.

Mindlin, Sergio E. and Howard Aldrich (1975), "Interorganizational Dependence: A Review of the Concept and a Reexamination of the Findings of the Aston Group," *Administrative Science Quarterly*, 20 (September), 382–92.

Moran, Peter (2005). "Structural vs. Relational Embeddedness: Social Capital and Managerial Performance," *Strategic Management Journal*, 26 (December), 1129–51.

Narayandas, Das and Kastuti Rangan (2004), "Building and Sustaining Buyer–Seller Relationships in Mature Industrial Markets," *Journal of Marketing*, 68 (July), 63–77.

North, Douglass Cecil (1990), *Institutions, Institutional Change, and Economic Performance*. New York: Norton.

Oliver, Christine (1991), "Strategic Responses to Institutional Processes," *Academy of Management Review*, 16 (January), 145–79.

Ostrom, Elinor (2000), "Collective Action and the Evolution of Social Norms," *Journal of Economic Perspectives*, 14 (Summer), 137–58.

Payan, Janice M. and Richard G. McFarland (2005), "Decomposing Influence Strategies: Argument Structure and Dependence as Determinants of the Effectiveness of Influence Strategies in Gaining Channel Member Compliance," *Journal of Marketing*, 69 (July), 66–79.

Perrow, Charles (1974), "Is Business Really Changing?" *Organizational Dynamics*, 2 (Summer), 31–44.

Peter, J. Paul, Gilbert A. Churchill Jr., and Tom J. Brown (1993), "Caution in the Use of Difference Scores in Consumer Research," *Journal of Consumer Research*, 19 (4), 655–76.

Pfeffer, Jeffrey and Gerald R. Salancik (1978), *The External Control of Organizations: A Resource-Dependence Perspective*. New York: Harper & Row.

Podsakoff, Philip M., Scott B. MacKenzie, Jeong-Yeon Lee, and Nathan P. Podsakoff (2003), "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, 88 (5), 879–903.

Romanelli, Elaine and Olga M. Khessina (2005), "Regional Industrial Identity: Cluster Configurations and Economic Development," *Organizational Science*, 16 (4), 344–58.

Schroeder, Roger G., Kimberly A. Bates, and Mikko A. Junttila (2002), "A Resource-Based View of Manufacturing Strategy and the Relationship to Manufacturing Performance," *Strategic Management Journal*, 23 (2), 105–117.

Scott, W. Richard (1983), "The Organization of Environments: Network, Cultural, and Historical Elements," in *Organizational Environments: Ritual and Rationality*, John W. Meyer and W. Richard Scott, eds. Beverly Hills, CA: Sage Publications, 155–75.

——— (1995), *Institutions and Organizations*. Thousand Oaks, CA: Sage Publications.

——— (2001), *Institutions and Organizations*, 2d ed. Thousand Oaks, CA: Sage Publications.

Smith, J. Brock and Donald W. Barclay (1997), "The Effects of Organizational Differences and Trust on the Effectiveness of Selling Partner Relationships," *Journal of Marketing*, 61 (January), 3–21.

Srinivasan, Raji, Gary L. Lilien, and Arvind Rangaswamy (2002), "Technological Opportunism and Radical Technology Adoption: An Application to E-Business," *Journal of Marketing*, 66 (July), 47–60.

Steenkamp, Jan-Benedict E.M. and Hans Baumgartner (1998), "Assessing Measurement Invariance in Cross-National Consumer Research," *Journal of Consumer Research*, 25 (June), 78–90.

Strang, David and John W. Meyer (1994), "Institutional Conditions for Diffusion," in *Institutional Environments and Organizations*, W. Richard Scott and John W. Meyer, eds. Thousand Oaks, CA: Sage Publications, 100–112.

Su, Chenting, Edward F. Fern, and Keying Ye (2003), "A Temporal Dynamic Model of Spousal Family Purchase-Decision Behavior," *Journal of Marketing Research*, 40 (August), 268–81.

Suchman, Mark C. (1995), "Managing Legitimacy: Strategic and Institutional Approaches," *Academy of Management Review*, 20 (3), 571–610.

Ulaga, Wolfgang and Andreas Eggert (2006), "Value-Based Differentiation in Business Relationships: Gaining and Sustaining Key Supplier Status," *Journal of Marketing*, 70 (January), 119–36.

Ulrich, David and Jay B. Barney (1984), "Perspectives in Organizations: Resource Dependence, Efficiency, and Population," *Academy of Management Review*, 9 (3), 471–81.

Wathne, Kenneth H. and Jan B. Heide (2004), "Relationship Governance in a Supply Chain Network," *Journal of Marketing*, 68 (January), 73–89.

White, J. Chris, P. Rajan Varadarajan, and Peter A. Dacin (2003), "Market Situation Interpretation and Response: The Role of Cognitive Style, Organizational Culture, and Information Use," *Journal of Marketing*, 67 (July), 63–79.

Wold, Herman (1985), "Systems Analysis by Partial Least Squares," in *Measuring the Unmeasurable*, P. Nijkamp, L. Leitner, and N. Wrigley, eds. Martinus Nijhoff: Dordrecht.

Zahay, Debra and Abbie Griffin (2004), "Customer Learning Processes, Strategy Selection, and Performance in Business-to-Business Service Firms," *Decision Sciences*, 35 (2), 169–203.

Zucker, Lynne G. (1977), "The Role of Institutionalization in Cultural Persistence," *American Sociological Review*, 42 (October), 726–43.

Eric (Er) Fang, Robert W. Palmatier, Lisa K. Scheer, & Ning Li

Trust at Different Organizational Levels

The authors explore the effects of trust at three distinct organizational levels in a marketing collaboration: interorganizational trust between collaborating firms, each firm's agency trust in its own representatives assigned to a collaborative entity (coentity), and intraentity trust among the representatives assigned to the coentity. Dyadic survey and longitudinal objective performance data from 114 international joint ventures indicate that trust at each level has unique effects but similarly influences the collaborating firms' resource investments or the coentity's use of those resources. Interorganizational and agency trust motivate resource investments in the coentity, particularly in the context of a differentiation strategy, whereas intraentity trust promotes coordination within the coentity, and interorganizational trust and a differentiation strategy magnify that effect. Intraentity trust can also undermine coentity responsiveness to environmental change, especially when joined by interorganizational trust between collaborating firms and formalized decision making within the coentity. These findings demonstrate that managing and building trust at multiple levels is critical to the success of interorganizational marketing collaborations.

Keywords: trust, interorganizational relationship, strategy, performance

Interorganizational collaboration is critical to marketers' success in the global marketplace; even the few firms that have the necessary resources to consider operating independently typically choose to focus on their distinctive competencies and collaborate with partners for more peripheral operations. Interfirm marketing collaborations often entail the formation of a collaborative entity or coentity—an identified set of employees from each collaborating firm tasked to work together to achieve collaborative outcomes. Firms form coentities for diverse reasons, such as to develop new products (Rindfleisch and Moorman 2001), strengthen supply chains (Wathne and Heide 2004), reduce operating costs (Cannon and Homburg 2001), reach new markets (Bamford, Ernst, and Fubini 2004), devise industry standards (e.g., the 3G project of Qualcom, Ericsson, and others), or serve specific customers. A coentity may be as simple as an informal work group or as elaborate and complex as an equity-based joint venture.

Firms form collaborative entities to generate value and achieve objectives that would be difficult, if not impossible, to achieve independently, but collaborations often fail to reach these goals. A recent study of senior executives suggests that "often the ... difficulties within a venture are poor relations" (Buchel 2003, p. 91). Buchel (2003) describes how the fundamental lack of trust among key constituents led to conflict and the ultimate dissolution of a venture between Leica and Zeiss that aimed to become the key supplier of electron microscopes. Coentity failures often are attributed to a lack of trust (Inkpen and Beamish 1997), in which context trust refers to a willingness to rely on another (Moorman, Zaltman, and Deshpandé 1992), the belief that a partner will meet its future obligations, and the confidence that the other will behave with integrity and benevolence rather than pure self-interest (Scheer and Stern 1992). Trust is especially critical in marketing collaborations, which often involve large, specific investments that could prompt opportunistic resource transfers, involve multiple constituents with diverging and/or conflicting objectives, and require a high level of cooperation to generate value (Das and Teng 1998; Rindfleisch and Moorman 2001). This unique context of multiple firms, diverse motives, and mixed loyalties makes the role of trust in collaborative entities particularly complex. As Zaheer, McEvily, and Perrone (1998, p. 141) note, "considerable ambiguity is evident in the literature about the precise role of trust as it operates at different levels of analysis and its influence on performance."

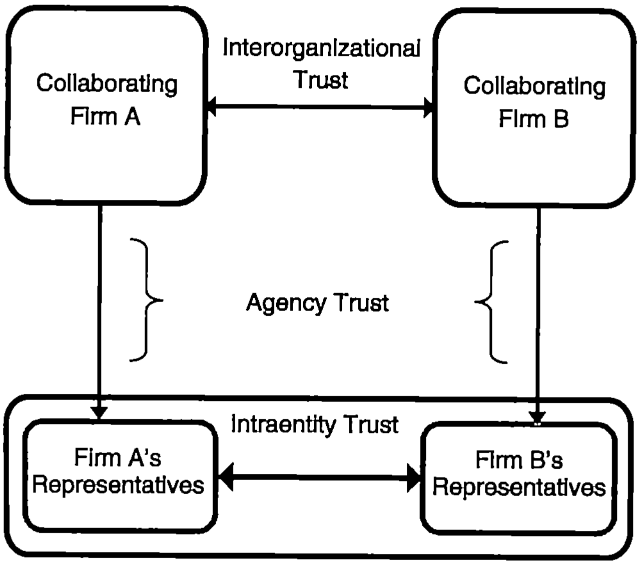
Any marketing collaboration that creates a formal or informal coentity depends on trust at three distinct levels: interorganizational trust, or the collaborating firms' mutual trust; agency trust, or each firm's trust in its own representatives assigned to the coentity; and the intraentity trust between the collaborating firms' representatives within the coentity (see Figure 1). Although each of these trust levels may influence coentity performance through different pathways with potentially differential effects, "issues of similarities and differences in collaborative trust at the person,

---

Eric (Er) Fang is an assistant professor, Lerner School of Business and Economics, University of Delaware (e-mail: fange@lerner.udel.edu). Robert W. Palmatier is Assistant Professor of Marketing and Evert McCabe Faculty Fellow, Michael G. Foster School of Business, University of Washington (e-mail: palmatrw@u.washington.edu). Lisa K. Scheer is Emma S. Hibbs Distinguished Professor and Associate Professor of Marketing, Robert J. Trulaske, Sr., College of Business, University of Missouri (e-mail: scheer@missouri.edu). Ning Li is an assistant professor, School of Management, George Mason University (e-mail: nli@gmu.edu).

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

FIGURE 1
Trust at Multiple Levels Within a Collaborative Entity (Coentity)



group, and firm levels have received limited empirical attention from scholars" (Currall and Inkpen 2002, p. 479).

We shed light on this research gap by examining how the three levels of trust affect coentity performance and how coentity characteristics may moderate these effects. Our research extends marketing theory by integrating previous unilevel research on trust to provide a more holistic picture of its complex interplay at multiple organizational levels. Teasing apart the differential roles of the various levels of trust, rather than using the more typical single-level perspective, is a critical step to reducing the high failure rates of interfirm collaborations (Inkpen and Beamish 1997).

Furthermore, previous studies have tended to examine the effects of trust in isolation without considering other relevant coentity characteristics, such as decision making and strategic focus. As Schoorman, Mayer, and Davis (2007, p. 351) articulate, the existing studies "have neglected many specific context variables that would be relevant to trust," and the failure to address these factors may lead to an incomplete or even incorrect picture of the impacts on firm performance. This bias further distorts existing knowledge when combined with the assumption in current literature that building trust at any level leads to desirable outcomes (Dirks and Ferrin 2001; Morgan and Hunt 1994). We address this issue by exploring the moderating effect of coentity characteristics, as well as both positive and negative outcomes of different levels of trust, and therefore provide a more nuanced view of how a coentity can better manage different levels of trust to achieve improved financial performance.

Drawing on dyadic survey data and longitudinal objective performance data from 114 collaborative entities, we find that trust at diverse levels within a marketing collaboration differentially affects the success of the collaboration and interacts with trust at other levels, as well as management decision processes and the coentity's strategic focus, to both enhance and undermine coentity financial performance. Although we examine a specific type of coentity, we contend that the theoretical model, principles, and underlying relationships among constructs generalize to a wide variety of marketing collaborations, such as joint ventures, supply chains, product development task forces, project teams, and work groups.

We first offer our theoretical model and research hypotheses. Then, we discuss our research context and methodology. After summarizing the results from our research, we discuss the theoretical and managerial implications of our findings, note several limitations, and highlight promising avenues for further research.

Theoretical Background

Trust at Multiple Levels in a Coentity

A coentity forms when at least two collaborating firms agree to contribute designated representatives to an identifiable group with an independent mission that benefits both firms. At one extreme, the coentity may be an informal intersection of the two firms, such as a product development or project management team created by identifying specific members who remain formally attached to their original firms but simultaneously interact with their counterparts at the other firm. At the other extreme, the coentity may be a joint venture or wholly owned independent company that is distinct from either collaborating firm but staffed initially with personnel supplied by the firms.

However formal or informal the configuration of the coentity, each collaborating firm contributes tangible and intangible resources with the intention of achieving common, stated objectives, as well as any additional individual, private objectives each may have for the coentity. The firms' representatives assigned to the coentity must work together to integrate those resources if the coentity is to achieve its stated objectives. As boundary spanners, these representatives face a mixed-motive situation: As agents of the firm that assigned them to the coentity, they must represent the interests of their own collaborating firm, but to ensure the coentity's success, they must become colleagues and form strong working relationships with their counterparts who represent the other firm. Employees assigned to a joint new product development team, for example, have the responsibility to contribute to the successful development of a product, but they may also be required to gather specific knowledge and technology possessed by the other collaborating firm for the benefit of their initial firm.

Collaborative entities are particularly complex because relationships between parties exist simultaneously at multiple levels, and the coentity itself owes allegiance to multiple constituents with potentially conflicting objectives. We contend that trust at three distinct levels is important to the success of any marketing collaboration that involves the creation of a coentity (see Figure 1): interorganizational trust between the collaborating firms, each firm's agency trust in its representatives assigned to the coentity, and intraentity trust among those representatives. In Table 1, we list research on trust at these levels that offers insight into their effects on coentity performance. However, most of these studies focus on trust only at a single level, raising the possibility that the type of trust they explicitly examine may be confounded with trust at another, unobserved level. Because trust is critical to the survival and long-term performance of collaborative entities (Dirks and Ferrin 2001; Doney and Cannon 1997; Morgan and Hunt 1994), we believe that disentangling the role of trust at multiple levels is critical.

How does trust at each level generate actions that ultimately drive performance? The resource-based view (RBV) highlights two mechanisms through which an organization can generate superior gains and competitive advantages (Barney 1991), namely, resource investments and resource utilization (Dyer and Singh 1998; Morgan, Kaleka, and Katsikeas 2004; Palmatier, Dant, and Grewal 2007). In collaborative entities, the interplay between these two mechanisms becomes complicated because each collaborating firm invests resources but lacks direct control over their usage. This disconnect between control over resource investment versus utilization, coupled with the conflicting loyalties of individual members assigned to the coentity, suggests that both investment and utilization can be affected by trust between the parties involved in the collaboration. We first examine how trust at multiple levels may affect resource investment and utilization and then turn our attention to how these processes may affect coentity performance (see Figure 2).

Effects of Trust at Multiple Levels on Collaborating Firms' Resource Investments

Collaborating firms' resource investments in a coentity typically include both nonfungible physical assets, such as manufacturing facilities or specialized machine tools, and human assets, such as employees who possess irreplaceable tacit knowledge. Because these specific investments would be lost if the coentity were dissolved prematurely, they create a lock-in situation (Williamson 1985) that exposes each collaborating firm to the other's opportunism, which can lead to suspicion and conflict (Bamford, Ernst, and Fubini 2004). Investment in a coentity carries considerable risk because the involvement of multiple parties with diverse goals and partially conflicting interests increases the potential leakage of the firms' tacit knowledge and strategic resources. We theorize that trust between the collaborating firms affects their resource investments, as does each collaborating firm's agency trust in its own representatives assigned to that coentity.

Trust in interorganizational relationships increases relationship investments, communication, and performance and reduces costs and opportunistic behaviors (Selnes and Sallis 2003; Smith and Barclay 1997). In the absence of trust, conflict between collaborating firms may prevent future investments or even lead to the withdrawal of existing investments (Inkpen and Beamish 1997). Mutual trust functions as a safeguarding and controlling mechanism that promotes information sharing and reduces collaborating firms' incentives and propensity to engage in opportunistic behaviors (Lane, Salk, and Lyles 2001). We hypothesize that interorganizational trust between collaborating firms reduces the perceived risk of opportunism and conflict, leading to a greater willingness to invest resources in the coentity.

> H_1: Interorganizational trust between collaborating firms positively affects each firm's resource investment in the coentity.

The managers assigned to the coentity develop situation-specific knowledge and cultivate unique interpersonal relationships that make them critical to the coentity's success. They also experience mixed motives and conflicting loyalties between the collaborating firm they represent and their colleagues within the coentity (Bonoma and Johnston 1978; Zaltman and Bonoma 1977). Therefore, a collaborating firm confronts potential opportunistic behavior by representatives who may strive to negotiate better personal compensation or positions (Wathne and Heide 2000), shift their primary allegiance to the coentity, or misuse knowledge generated within the coentity to compete with the collaborating firm. Kumar and Seth (1998) suggest that staffing the coentity with trusted representatives encourages collaborating firms to provide the ongoing resources needed for the coentity's success. That is, a collaborating firm is more likely to share additional intellectual property, manufacturing technology, and other sensitive resources with its coentity when its trust in its own representatives is greater. As a firm's trust in its representatives increases, it becomes more committed to those representatives, which motivates the firm to increase support (Child and Mollering 2003; Gilliland and Bello 2002). When a collaborating firm's agency trust is greater, the firm perceives less risk, expects invested resources to be used wisely, and is willing to make greater resource investments in the coentity. Thus:

> H_2: A collaborating firm's agency trust in its representatives within a coentity positively affects its resource investment in the coentity.

Effects of Trust at Multiple Levels on Coentity Resource Utilization

Resource utilization involves the idiosyncratic organizational processes used to combine, integrate, and redeploy component resources while responding to changing environmental conditions (Moran and Ghoshal 1999; Reed and DeFillippi 1990). We examine one internally focused process, coordination, and one externally focused process, responsiveness (Eisenhardt and Martin 2000; Teece, Pisano, and Shuen 1997). Coentity coordination is the representatives' effectiveness in working together to integrate, combine, and deploy resources. Coentity responsiveness is the coentity's effectiveness in marshalling and redeploying resources in response to environmental changes. We theorize that both intraentity trust and interorganizational trust between collaborating firms affect resource utilization.

Collaborating firms control their resource investments, but the use of these resources depends on the people directly involved in the day-to-day operations of the coentity. The interaction of key representatives assigned to a coentity involves "a continuous negotiation of the overarching interests" of their parent firms (Loess and Yavas 2003,

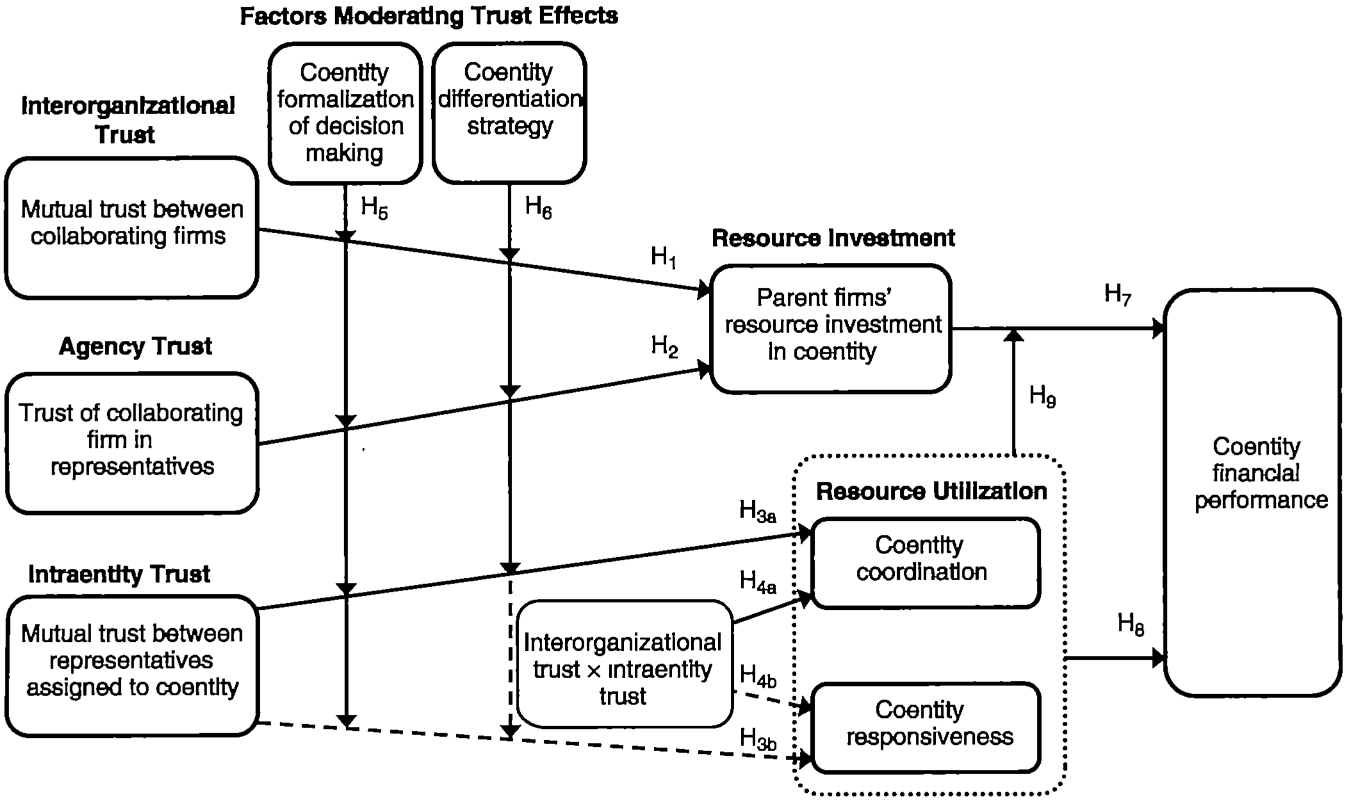
TABLE 1
Summary of Illustrative Research on the Influence of Interorganizational, Agency, and Intraentity Trust on Key Outcomes

Illustrative Research	Context	Moderator Variables	Outcome Variables	Key Findings
Interorganizational Trust Between Two Organizations				
Grayson and Ambler (1999)	Advertising agencies and their clients	Relationship age	Commitment, quality of interaction, involvement, and performance	Trust positively affects commitment, interaction quality, involvement, and performance. Trust's effect on performance is lower for long-term versus short-term relationships.
Selnes and Sallis (2003)	Suppliers to medium-sized and large business customers	Relationship learning	Multidimensional measure of relationship performance	Trust positively affects relationship performance (payoff of investments in relationship, synergies, flexibility, lower costs). There is a negative interaction effect of trust and relationship learning on performance, leading to less information exchange, fewer meetings, less evaluation of relationship, and less adjusting to end-user performance.
Smith and Barclay (1997)	Suppliers of computer equipment to business customers		Relationship investments, communication, and opportunism	Mutual perceived trustworthiness positively affects relationship investments, communication, and performance and reduces opportunism. Relationship investment and communication have a subsequent positive effect on performance.
Zaheer, McEvily, and Perrone (1998)	Electrical equipment manufacturers and their component suppliers		Negotiation cost, conflict, and performance	Trust directly affects performance, but the effects are not mediated through reduced conflict or reduced negotiation costs. According to post hoc analysis, trust's effect on performance may be mediated not by gains in efficiencies as much as by "exchange of personnel," shared decision making, and improved coordination (p. 155).
Agency Trust in an Organization's Representatives				
Child and Mollering (2003)	Corporate managers and their representatives in foreign (Chinese) ventures		Growth in sales and profits	Corporate managers' trust in their foreign entities' performance is positively related to growth in sales and profits.

TABLE 1
Continued

Illustrative Research	Context	Moderator Variables	Outcome Variables	Key Findings
Gilliland and Bello (2002)	Suppliers and their industrial product distributors and dealers		Loyalty commitment	Suppliers' trust in their distributors positively affects their loyalty commitment to the relationship.
Intraentity Trust Between Two Groups Within One Organizational Entity				
Atuahene-Gima and Li (2002)	Sales managers and salespeople within one organization	Output control, supervisor accessibility, achievement orientation, and role ambiguity	Sales performance	Trust is likely to enhance entity sales performance when output control is high (low) for the Chinese (U.S.) sample, supervisor accessibility is low, achievement orientation is high, and, in the Chinese sample, role ambiguity is high.
Dirks and Ferrin (2001)	Review of 40 years of research on intraorganizational trust	Goals, motives, incentives, and other organizational factors	Communication, citizenship behaviors, conflict, individual and entity performance, and satisfaction	Review of literature suggests that trust has a strong effect on organizational citizenship behaviors, satisfaction, and individual performance, but trust's effect on group behaviors and performance is mixed. The authors argue that trust may provide the organizational context for other drivers of performance to operate; trust may interact with employees' goals, motives, and other organizational factors.
Gargiulo and Benassi (2000)	Special unit responsible for managing various change initiatives within an organization		Coordination failures	The results support structural hole theory; cohesive social bonds undermine the ability of social actors to adapt to changes in the environment. For example, the autonomy of managers (e.g., communication levels) in dealing with key third parties is reduced for managers with strong relationships (e.g., trust) with others in the group.
Maltz and Kohli (1996)	Different functional departments within one organization		Frequency, accuracy, clarity, and relevance of information	Trust in another department positively affects the frequency, accuracy, clarity, and relevance of information provided by that department.
Tsai and Ghoshal (1998)	Different business units within one organization		Resource exchange and resource combination	Trust in other business units positively affects resource exchange and resource combination.

FIGURE 2
Impact of Trust at Different Levels on Collaborative Entity (Coentity) Financial Performance


Factors Moderating Trust Effects
Coentity formalization of decision making
Coentity differentiation strategy
Interorganizational Trust
Mutual trust between collaborating firms
H_5
H_6
H_1
Resource Investment
Parent firms' resource investment in coentity
H_7
Agency Trust
Trust of collaborating firm in representatives
H_2
H_9
Coentity financial performance
Resource Utilization
H_{3a}
Coentity coordination
Intraentity Trust
H_{4a}
Mutual trust between representatives assigned to coentity
Interorganizational trust × intraentity trust
H_{4b}
H_8
Coentity responsiveness
H_{3b}

Notes: Dashed lines represent hypothesized negative effects.

p. 313), which can be problematic if insufficient trust marks their relationships. Representatives from one firm often have difficulty understanding the culture, knowledge, and competency of their colleagues from the collaborating firm, a situation that may be exacerbated by cultural distance. Information asymmetry may lead a firm's representatives to resist adapting their work patterns to those of their counterparts (Jeffries and Reed 2000) or to free ride on the efforts of the other firm's representatives (Wathne and Heide 2000). In contrast, intraentity trust among colleagues increases the frequency and accuracy of information exchange and resource coordination (Dirks and Ferrin 2001; Maltz and Kohli 1996), motivates cooperative decision making, reduces fears of exploitation, and increases resource sharing and risk taking (Chiles and McMackin 1996). Therefore, we posit that greater intraentity trust among collaborating firms' representatives increases coordination within the coentity.

In contrast, intraentity trust could reduce the coentity's responsiveness to external conditions. Drawing on structural hole theory, Gargiulo and Benassi (2000) demonstrate that social capital has a dark side; strong relational bonds can filter external information and generate a cognitive lock-in that isolates bonded parties from the outside world (Grabher 1993). People with strong bonds tend to build self-reinforcing business processes that make them less able to adapt to environmental changes (Tushman and Anderson 1986). When representatives within the coentity build interpersonal relationships and trust one another, routine rigidity is more likely to develop because they feel less need to adjust established procedures or to monitor their counterparts.[1] Therefore, as intraentity trust increases, collaborative entities may develop relational inertia and routine rigidity, which reduces the coentity's responsiveness by minimizing its motivation to engage in environmental scanning and impeding its ability to redeploy resources quickly in the face of environmental change. Thus:

> H_3: Intraentity trust between representatives of both collaborating firms (a) positively affects coentity coordination and (b) negatively affects coentity responsiveness.

Confidence in another's trustworthiness provides a motive to behave in a trusting manner, but whether that motivation is manifest in actions depends on other factors. On the basis of a review of 40 years of research in management literature, Dirks and Ferrin (2001) posit that intraorganizational trust may interact with other motives and objectives to determine the level of cooperation. Researchers investigating channels and interorganizational relationships in marketing and management (e.g., Doney and Cannon 1997; Zaheer, McEvily, and Perrone 1998) note the importance of the interrelationships among constructs at both the interfirm (e.g., supplier and customer firms) and interpersonal (e.g., salesperson and buyer) levels to comprehend the full effects of cross-firm ties. Extending this logic, we consider how the interfirm relationship between collaborating firms could affect relationships among their representatives who must work together within the coentity.

---

[1] We thank a reviewer for this insight.

We posit that the extent to which intraentity trust affects coordination and responsiveness depends on the interorganizational trust between the collaborating firms. Regardless of a coentity manager's confidence in a counterpart's trustworthiness, acting on that trust is inherently risky if it runs counter to the collaborating firm's beliefs. Representatives face significantly less risk in acting on their own trust in their counterparts if interorganizational trust between the collaborating firms reinforces or confirms their intraentity trust. Thus, we anticipate that interorganizational and intraentity trust have an interactive effect on coentity coordination and responsiveness.

> H_4: Collaborating firms' interorganizational trust amplifies the (a) positive effect of intraentity trust on coentity coordination and (b) negative effect of intraentity trust on coentity responsiveness.

Other Factors Moderating the Effects of Trust

Trust provides the motivation to act, but other elements may encourage or discourage a person from actually translating that motivation into behavior. Consistent with our desire to develop a generalized model of marketing collaborations, we investigate two factors that are relevant for a wide variety of collaborative entities: the level of formalization in the coentity's decision making and the degree to which the coentity implements a differentiation strategy.

Formalization of decision making refers to the extent to which the decision-making process emphasizes and follows specific rules and procedures (Zaltman, Duncan, and Holbek 1973). Formalization enhances the likelihood that a trustee behaves cooperatively and not opportunistically (Das and Teng 1998), provides the trustor with greater confidence that his or her trust will not be abused, and promotes more risk-taking behaviors (Kogut 1989). Thus, we theorize that formalized decision making within a coentity enhances the entire constellation of risk-taking behaviors associated with each type of trust.

Specifically, we argue that formalized decision making enhances the positive effects of the collaborating firms' interorganizational trust and agency trust on resource investment and magnifies the positive effect of intraentity trust on coentity coordination. We similarly anticipate that formalized decision making inflates the negative effect of intraentity trust on coentity responsiveness. When decision-making policies and procedures become more formalized, each party's duties become increasingly predefined and divided (Kelly and Amburgey 1991). Routine rigidity and relational inertia bred by intraentity trust increase as decision making becomes more formalized, which further reduces the coentity's adaptability and responsiveness to change. Thus:

> H_5: More formalized decision making within the coentity amplifies (a) interorganizational trust's positive effect on resource investment, (b) agency trust's positive effect on resource investment, (c) intraentity trust's positive effect on coordination, and (d) intraentity trust's negative effect on responsiveness.

A differentiation strategy captures the strategic emphasis of the coentity in creating and delivering unique customer benefits in new and distinct ways (Porter 1985), an effort that often requires highly specialized investments (Ghosh and John 1999). Because these specialized investments likely are context and market specific, a collaborating firm's vulnerability to opportunism increases. Thus, when coentities focus more on a differentiation strategy, trust becomes an even more critical safeguard against opportunism, and the positive effects of interorganizational and agency trust on resource investment should be greater.

A differentiation strategy also requires constant adjustments to provide unique value to customers. For example, state-of-the-art technology providers must incorporate frequent engineering and technology changes into their product offerings. Because a differentiation strategy creates adaptation and coordination challenges for the managers assigned to a coentity, opportunities for conflict increase; intraentity trust becomes more critical to avoid and resolve such conflicts while maintaining the high level of coordination needed to facilitate strategic changes. Thus, we expect that intraentity trust has a more positive impact on coordination when coentities adopt a differentiation strategy.

In contrast, a differentiation strategy may lessen intraentity trust's negative impact on responsiveness. The adaptation requirements inherent in pursuing a differentiation strategy force the coentity's managers to develop new processes and form new relational ties. The high level of change required by a differentiation strategy also diminishes the potential relational inertia and routine rigidity associated with intraentity trust (Ghosh and John 1999). Furthermore, because differentiation generates frequent internal changes, role ambiguity and potential conflict emerge among representatives. Thus, the intricacies and demands of pursuing a differentiation strategy counter the tendency of intraentity trust to dampen responsiveness.

> H_6: A coentity's differentiation strategy (a) amplifies interorganizational trust's positive effect on resource investment, (b) amplifies agency trust's positive effect on resource investment, (c) amplifies intraentity trust's positive effect on coordination, and (d) suppresses intraentity trust's negative effect on responsiveness.

Effects of Resource Investment and Usage on Coentity Financial Performance

Consistent with the RBV, we posit that the collaborating firms' resource investments in the coentity positively affect the coentity's financial performance. Each collaborating firm invests tangible and intangible resources, such as specialized development tools, information technology, equipment, and employees' knowledge and skills, to implement specialized tasks. These resource investments form the coentity's component competencies, which generate competitive advantage (Prahalad and Hamel 1990), promote the achievement of objectives, and generate superior performance (Smith and Barclay 1997). Consistent with the RBV, we hypothesize that the coentity's resource utilization promotes financial performance; both coordination and responsiveness reflect dynamic capability and therefore should improve performance (Eisenhardt and Martin 2000).

Coordination and responsiveness also can leverage the effect of resource investments on performance (Makadok 2001). Collaborative entities operated by people who coordinate their actions and respond to external changes should generate higher returns on the resources at their disposal. Thus, we hypothesize that there is a positive interaction between the resource utilization processes and resource investment; specifically, coentity coordination and responsiveness should enhance the positive effect of the collaborating firms' resource investments on coentity performance.

H_7: Collaborating firms' resource investments in a coentity positively affect the coentity's financial performance.

H_8: A coentity's (a) coordination and (b) responsiveness positively affect its financial performance.

H_9: A coentity's (a) coordination and (b) responsiveness positively moderate the effects of collaborating firms' resource investments on the coentity's financial performance.

Research Methods

Research Context

We test our hypotheses using a sample of equity-based collaborative entities, specifically, international joint ventures (IJVs). Joint ventures provide an ideal context for testing our model because they consist of collaborating firms that invest tangible and intangible resources in a clearly defined coentity and are staffed by representatives of both firms who are tasked to pursue the objectives set by those firms. In addition, IJVs often have diverse degrees of decision-making formalization and differentiation strategies, which offers an opportunity to test how these factors may leverage the effects of trust on coentity performance. Our focus on IJVs in China offers additional advantages. Notably, lack of trust likely represents a critical factor in the failure of collaborative ventures in China (Luo 2000). Furthermore, China provides a managerially relevant context because equity-based IJVs are frequent vehicles for foreign market entry, and the Chinese market itself is economically important. By the end of 2005, approximately 320,000 collaborative ventures had been established in China; the associated $600 billion in foreign direct investments constitutes approximately half of such investments in developing countries. In 2004, China surpassed the United States to become the largest recipient of foreign direct investments, mainly through IJVs.

Survey Data Collection Process

Our sample consists of 200 IJVs in China, gleaned randomly from high-tech zones in Jiangsu Province. Jiangsu is second among Chinese provinces in luring foreign capital, third in generating gross domestic product, and representative of the nation's cultural norms and standards (Luo 2005). Moreover, IJVs in this region are representative of those in China (Luo 2005). Each venture involves at least one collaborating firm based outside China and at least one collaborating firm headquartered in China. When more than two collaborating firms are involved, we focus on the largest domestic and the largest foreign participants.

Within our specific research context, we adopt the terminology our study participants used: The coentity is a "joint venture," the collaborating firms are the "foreign parents" and the "local parents," and the collaborating firms' respective representatives to the coentity are "foreign partners" and "local partners." To test our theoretical model, we engaged in a two-stage data collection process. First, we collected matched, dyadic survey data from the senior managers assigned to the IJV by each parent. Theoretical considerations and field interviews guided the development of our measures and our survey design. The original English version of the survey was translated into Chinese and back translated into English to ensure equivalence. We pretested the questionnaire for instrument validity with paired dyads of 30 managers who represent 15 joint ventures. The results reveal a high degree of internal consistency in the responses to the questionnaire items between the paired managers (Guttman split-half $R > .86$).

In May 2003, we placed telephone calls to the general managers of the 200 IJVs in our initial sample to explain the purpose of the study, stress the endorsement of the relevant government administrative agency, and offer a summary copy of the aggregate results and customized analyses in return for their participation. We also verified that a clear distinction existed between each parent company and its representatives within the venture (i.e., the local parent is distinct from the local partner, and the foreign parent is distinct from the foreign partner). Then, we eliminated any ventures for which there was no such clear distinction.[2] Of the initial 200 joint ventures in our sample, 146 met our qualification standards and agreed to participate.

Surveys were personally delivered to the two senior managers in each IJV, one associated with the local partner and the other associated with the foreign partner. In the initial visit, the interviewer assured respondents of their confidentiality and discussed the government agency endorsement to enhance the researchers' credibility. Personal delivery establishes a tangible connection between the manager and the researcher, and the knowledge that the same researcher would personally return to pick up the survey increased the manager's obligation to participate. Even this minimal personal interaction appears to increase response among Chinese participants. We received completed questionnaires from representatives of both parent firms (i.e., the manager of the local partner and the manager of the foreign partner) for 131 joint ventures, for a 90% response rate.

We evaluated each informant's overall knowledge of the IJV's operation, strategy, resources, and capabilities, as well as his or her involvement in the strategy making and daily operations of the joint venture, using seven-point scales (1 = "very low," and 7 = "very high"). We dropped five cases for which an informant expressed less than 4 on either knowledge or involvement, which left 126 IJVs in our sample. The 252 informants average 6.3 on the knowledge and 5.9 on the involvement scales.

---

[2] In some IJVs, the Chinese parent firm becomes the local partner in the coentity; in such cases, no distinct representatives are assigned to the joint venture.

In the second stage, we gathered longitudinal, objective performance data two years later (June 2005) from each joint venture's archives. This two-year period should provide adequate time for the effects of resource investment and resource usage to become manifest in the venture's tangible outcomes. Because of management turnover or ownership changes, we could not obtain archival data for 12 ventures. Therefore, our multistage data collection generated a final data set of 114 IJVs with primary data from both senior representatives of the collaborating parent firms and archival financial performance data (57% of the initial sample, 78% of those receiving surveys). Our comparisons of the 114 participating IJVs with the 86 nonparticipating ventures on available demographic characteristics yielded no significant differences, suggesting that nonresponse bias is not a concern. The IJVs in our data set operate in the consumer electronics, computer hardware, electronic components, medical supplies, and industrial controls industries.

Measurement

The foreign partner in the joint venture is the parent firm's representative within the coentity, just as the local partner represents the Chinese parent. Two constructs in our model—agency trust and resource investment—are specific to each collaborating parent firm; the senior manager in the corresponding IJV partner represents the single informant for the parent firm's agency trust and resource investment. For all other constructs, which pertain to the parents' dyadic relationship or to the joint venture itself, both the local and foreign partner senior managers serve as dual informants. Both foreign and local senior managers are suitable informants because of their high involvement in the joint venture (5.8 and 6.0 on a seven-point scale, respectively). Following Van Bruggen, Lilien, and Kacker's (2002) suggestions, we adopt a confidence-based weighted mean to obtain construct scores. In addition, we asked informants about their confidence in each conceptually similar section of the questionnaire, which minimized their cognitive burden because it captures any variations in their confidence about different questions.

To validate that a parent's representative could provide useful data about that parent firm's agency trust and interorganizational trust, we contacted additional informants at 21 parent firms and compared their reports with those of their representatives. We evaluated interrater reliability between the representatives and the parent firm informants using intraclass correlation coefficient (ICC), the widely used indicator of interjudge reliability in behavioral science literature (McGraw and Wong 1996; Snedecor and Cochran 1980). The ICC indicates that the representatives reliably reported the parent firm's agency trust (.64, $p < .001$) and interorganizational trust (.70, $p < .001$). (For a summary of all measures, see the Appendix.)

Constructs specific to each collaborating parent firm. We measure each parent firm's agency trust using two items adapted from the work of Hewett and Bearden (2001) that pertain to the extent to which the parent firm assesses its representatives as reliable and serving its best interests. Using an eight-item formative scale adapted from the work of Jap and Ganesan (2000), we assess the resource investments in the joint venture by each parent firm after the coentity's formation; informants reported tangible and intangible resources invested in eight functional areas, such as product development, sales personnel and training, and research and development. Thus, we obtain separate measures for the local parent's agency trust and resource investment (from the local partner manager) and for the foreign parent's agency trust and resource investment (from the foreign partner manager).

Dyadic parent firm and coentity constructs. We develop a new scale to measure parent firms' interorganizational trust and use three items to assess the extent to which the collaborating firms consider their partner reliable and benevolent (ICC = .64, $p < .01$). For intraentity trust, we adapt three items from the work of Johnson and colleagues (1996) that assess the extent to which both parents' representatives in the coentity believe that their partners are reliable and benevolent (ICC = .70, $p < .001$). We measure coentity coordination, or the effectiveness with which the representatives work together to use invested resources (ICC = .72, $p < .001$), with four items adapted from the work of Jap (1999). We develop three new items to measure coentity responsiveness to rapid environmental changes (ICC = .67, $p < .001$). Finally, we measure formalization of decision making with a four-item scale adapted from the work of Jaworski and Kohli (1993) (ICC = .68, $p < .001$) and differentiation strategy using four items (ICC = .70, $p < .001$).

Control variables. We capture several control variables that theoretically could affect both parent firms' resource investments and the joint venture's resource utilization and financial performance. Dysfunctional competition is the degree to which the competitive behaviors of firms in the market are perceived as opportunistic or unfair; this measure is based on four items adapted from the work of Li and Atuahene-Gima (2001) (ICC = .71, $p < .001$). We also measure environmental dynamism, or the degree of change and unpredictability in the market environment, with four items adapted from the work of Jaworski and Kohli (1993) (ICC = .69, $p < .001$). We examine the variation across joint ventures according to several dimensions. Consistent with Luo (2005), we measure cultural distance between the foreign parent's country of origin and China using Kogut and Singh's (1988) index. We also determine the joint venture's years of operation. Similar to Rindfleisch (2000), if the parent firms' primary businesses are similar (i.e., competitors), we classify the coentity as a horizontal joint venture; if their primary businesses occupy different levels within the vertical channel, we classify it as a vertical joint venture. Finally, we control for whether the joint venture's primary industry is a business or consumer market.

We also account for various additional control mechanisms that may affect resource investment and/or utilization. First, we measure contract specificity, which could affect both parent firms' resource investments and the joint venture's resource utilization (Luo 2002), using three items adapted from the work of Jap and Ganesan (2000) (ICC = .65, $p < .001$). Second, we examine the potential impact of parent monitoring on resource investment using three items adapted from the work of John (1984). Third, we control for interest alignment between the partners, measured as the equity share difference between the joint venture partners (Kogut 1989), which could affect resource utilization.

Longitudinal Data Collection: Coentity Financial Performance

Two years after the dyadic survey data collection, we contacted each joint venture to obtain objective financial performance data, specifically, sales/total assets for 2005 and profits/investment for 2003–2005. On the basis of factor loadings obtained through confirmatory factor analysis, we aggregate these two indicators using weighted averages to account for their respective importance in contributing to the latent measure of financial performance (Gerbing and Anderson 1988).

Measurement Model Analysis

For each measurement scale for which dual informants provided data, ICC indicates that sufficient similarity exists to aggregate the responses. Aggregating multiple informants' responses has significant advantages for perceptual constructs, in that it avoids a reliance on the attributions of a single individual, a major source of common method bias (Ayers, Dahlstrom, and Skinner 1997). Using multiple informants also can enhance data accuracy significantly by reducing the random error associated with individual response data (Van Bruggen, Lilien, and Kacker 2002). Following Steenkamp and Baumgartner's (1998) recommendations, we first establish full metric invariance across the two informants. None of the measures differ significantly across the local and foreign representatives subsamples ($p > .05$), which gives us additional confidence in combining the data from the two sources.

We estimate three separate measurement models by grouping theoretically related constructs together into (1) trust constructs, (2) governance-related constructs (parent monitoring, contract specificity, and formalization of decision making), and (3) strategic behavior and environmental constructs (all others). We restrict each scale item's loading to its a priori specified factor and allow correlations among factors. The fit indexes for each model are good. For the trust constructs' measurement model, $\chi^2_{(29)} = 43.43$ ($p > .10$), comparative fit index (CFI) = .95, normed fit index (NFI) = .93, and root mean square error of approximation (RMSEA) = .05. For the governance constructs' measurement model, $\chi^2_{(24)} = 40.17$ ($p > .10$), CFI = .93, NFI = .92, and RMSEA = .05. For the strategic and environmental constructs' measurement model, $\chi^2_{(215)} = 401.02$ ($p > .10$), CFI = .91, NFI = .89, and RMSEA = .07. All factor loadings are positive and significant ($p < .01$). Composite reliabilities are greater than .70 (for details, see the Appendix).

The average variance extracted by each construct is greater than the square of the latent correlation between that construct and all other constructs in that measurement model (Fornell and Larcker 1981). In addition, we conduct pairwise chi-square difference tests for each pair of constructs in the overall model (Bagozzi, Yi, and Phillips 1991). Discriminant validity is supported if an unconstrained model demonstrates significantly better fit than a constrained model in which we constrain the correlation between those constructs to one ($\Delta\chi^2$ significant at $p < .01$). These analyses suggest that we achieved discriminant validity among the constructs in our study. We summarize the descriptive statistics and correlations in Table 2.

Hypothesis-Testing Procedure

We use two estimation procedures to test the hypotheses presented in our theoretical model. To test H_1–H_6, we estimate a series of four models and examine our resource investment and resource utilization variables using seemingly unrelated regression (SUR). Because the error terms associated with these variables could be correlated, SUR is appropriate; it enables the estimation of theoretically related sets of equations and permits error terms to be correlated across equations. In such conditions, SUR provides more efficient estimates than ordinary least squares (OLS) (Zellner 1962). To test H_7–H_9, which pertain to the antecedents of financial performance, we use OLS. Table 3 summarizes the results from our SUR analyses of foreign parent resource investment, local parent resource investment, coordination, and responsiveness, and Table 4 offers the results from our OLS analysis. After mean-centering the independent and moderating variables to reduce potential multicollinearity (Aiken and West 1991), we find that the variance inflation factors (ranging from 1.24 to 3.11) suggest no serious problems.

Results

Main Effects of Trust on Resource Investment and Utilization

Trust at each level has a direct effect on resource investment or utilization, but we advise caution in interpreting these main effects because of the presence of significant interactions. Parent firms' interorganizational trust positively affects resource investments in the coentity (foreign: $\beta = .16$, $p < .05$; local: $\beta = .11$, not significant [n.s.]), but we find support for H_1 only among the foreign parent firms in our sample. Parent firm agency trust generates resource investments among both foreign ($\beta = .15$, $p < .10$) and local ($\beta = .27$, $p < .01$) parent firms, in support of H_2 (Table 3, Models 1 and 2). Intraentity trust between representatives positively affects coentity coordination ($\beta = .21$, $p < .05$) and negatively affects responsiveness ($\beta = -.27$, $p < .01$), in support of both H_{3a} and H_{3b} (Table 3, Models 3 and 4).[3]

---

[3] Among the control variables investigated, resource investments appear to decrease with dysfunctional competition (foreign: $\beta = -.11$, $p < .10$; local: $\beta = -.12$, $p < .10$) and increase with years of operation (foreign: $\beta = .16$, $p < .05$; local: $\beta = .13$, $p < .10$), as well as in business versus consumer markets (foreign: $\beta = .23$, $p < .05$; local: $\beta = .18$, $p < .05$). The foreign parent's monitoring positively affects its own resource investments ($\beta = .16$, $p < .05$). Interest alignment increases coordination ($\beta = .15$, $p < .05$) and respon-

TABLE 2
Means, Standard Deviations, and Correlations

			Correlation Matrix																		
Constructs	**M**	**SD**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
1. Interorganizational trust: mutual trust between collaborating firms	3.79	1.01	.62																		
2. Agency trust: collaborating firm's trust in its representatives assigned to the coentity (foreign parent)	3.97	1.21	.11	.60																	
3. Agency trust: collaborating firm's trust in its representatives assigned to the coentity (local parent)	4.12	1.14	.13	.14	.62																
4. Intraentity trust: mutual trust between representatives of different firms assigned to the coentity	4.35	1.20	.23*	.17	.08	.52															
5. Collaborating firm's resource investment (foreign parent)	3.65	1.11	.13	.22*	.19*	.09	N.A.														
6. Collaborating firm's resource investment (local parent)	3.81	1.02	.14	.05	.20*	.08	.16	N.A.													
7. Coentity coordination	4.15	1.07	.14	.12	−.12	.23*	.30*	.17	.69												
8. Coentity responsiveness	4.87	1.21	.16	.10	.10	−.25*	.19*	.12	.19*	.65											
9. Coentity financial performance	.17	.15	.04	−.15	.17	.24*	.29*	.21*	.08	.24*	N.A.										
10 Coentity formalization of decision making	3.87	1.27	.18*	.03	.04	−.12	.04	.17	.09	−.01	−.13	.66									
11. Coentity differentiation strategy	3.77	1.10	.14	.05	.05	.11	−.15	−.13	.08	−.09	.23*	−.11	.60								
12. Dysfunctional competition	4.50	1.35	−.24*	−.17	−.12	−.34*	−.31*	−.29*	−.08	.11	−.15	.09	.11	.59							
13. Environmental dynamism	4.32	1.03	−.25*	−.03	.04	.02	.11	.27*	−.13	.09	.09	−.07	.12	.32*	.58						
14. Cultural distance	5.73	1.45	−.04	.02	.13	.07	−.06	−.18*	.09	−.06	.19*	.04	−.03	.09	.12	N.A.					
15. Years of operation	5.52	2.22	.16	.07	.03	.13	.10	.03	.21*	.20*	−.10	−.11	.01	.10	.11	.11	N.A.				
16. Parent monitoring (foreign parent)	5.03	1.04	−.11	−.22*	.05	.03	.25*	−.05	−.10	.07	−.14	.17	.01	.01	.16	−.02	−.18*	.60			
17. Parent monitoring (local parent)	4.23	1.21	−.16	.06	−.14	−.06	.09	.14	.04	.15	.13	.20*	.04	.07	.08	−.05	−.20*	.11	.63		
18. Contract specificity	4.05	1.18	−.13	.05	.09	−.10	.14	.16	.05	.07	−.06	.18*	−.03	.10	.05	.13	−.14	.08	.17	.53	
19. Interest alignment	−.13	.25	.07	−.04	−.03	.14	.06	.04	.24*	.20*	−.01	−.05	.18*	.04	.13	−.04	.11	−.05	−.06	.02	N.A.

*$p < .05$.

Notes: Variance extracted reported on the diagonal for all latent constructs. N.A. = not applicable.

TABLE 3
Determinants of Collaborative Entity (Coentity) Resource Investment and Utilization

		Resource Investment (Foreign Parent)[a]	Resource Investment (Local Parent)[a]	Coentity Coordination[a]	Coentity Responsiveness[a]
Variables	**Hypotheses**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Control Variables					
Dysfunctional competition		–.11*	–.12*	–.03	–.09
Environmental dynamism		.09	.04	.06	–.02
Cultural distance		.04	.02	.07	.02
Years of operation		.16**	.13*	.07	.11
Contract specificity		.05	.06	.09	–.09
Joint venture type (1 = vertical; 0 = horizontal)		.06	.12	.16**	–.05
Industry (1 = business market; 0 = consumer market)		.23**	.18**	.02	.05
Interest alignment				.15**	.14*
Parent monitoring (foreign parent)		.16**			
Parent monitoring (local parent)			.11		
Main Effects					
Interorganizational trust	H_1	.16**	.11	.04	.05
Agency trust (foreign parent)	H_2	.15*			
Agency trust (local parent)	H_2		.27***		
Intraentity trust	H_{3a}, H_{3b}			.21**	–.27***
Coentity formalization of decision making		.03	.08	.14*	–.10
Coentity differentiation strategy		.08	.04	.15*	.17*
Moderating Effects					
Interorganizational trust × intraentity trust	H_{4a}, H_{4b}			.19**	–.22**
Formalization of decision making × interorganizational trust	H_{5a}	.05	.09		
Formalization of decision making × agency trust (foreign parent)	H_{5b}	.06			
Formalization of decision making × agency trust (local parent)	H_{5b}		.23**		
Formalization of decision making × intraentity trust	H_{5c}, H_{5d}			.13	–.29***
Differentiation strategy × interorganizational trust	H_{6a}	.21**	.13		
Differentiation strategy × agency trust (foreign parent)	H_{6b}	.23**			
Differentiation strategy × agency trust (local parent)	H_{6b}		.11		
Differentiation strategy × intraentity trust	H_{6c}, H_{6d}			.25**	.05
Weighted R^2		.27	.18	.20	.25

*$p < .10$.
**$p < .05$.
***$p < .01$.
[a] Standardized coefficients reported for SUR analyses.

Factors Moderating the Effects of Trust

Parent firms' interorganizational trust amplifies the positive effect of intraentity trust on coordination ($\beta = .19$, $p < .05$) and its negative effect on responsiveness ($\beta = -.22$, $p < .05$). Thus, we find support for both H_{4a} and H_{4b} (Table 3, Models 3 and 4). Formalization of decision making enhances both the positive effect of the local parent firm's agency trust on resource investment ($\beta = .23$, $p < .05$) and the negative effect of intraentity trust on responsiveness ($\beta = -.29$, $p < .01$). Therefore, H_{5b} receives partial support, and H_{5d} receives full support, but we must reject H_{5a} and H_{5c}.

---

siveness ($\beta = .14$, $p < .10$). Coordination is greater in vertical than horizontal joint ventures ($\beta = .16$, $p < .05$).

The coentity's differentiation strategy positively moderates the positive effects of interorganizational trust ($\beta = .21$, $p < .05$) and agency trust ($\beta = .23$, $p < .05$) on the foreign parent firms' resource investment. For the local parents, differentiation strategy does not significantly moderate the effect on investment of either interorganizational trust ($\beta = .13$) or agency trust ($\beta = .11$). That is, we find support for H_{6a} and H_{6b} only among foreign parent firms. A differentiation strategy enhances intraentity trust's positive effect on coordination ($\beta = .25$, $p < .05$) but does not significantly affect coentity responsiveness ($\beta = .05$, n.s.), in support of H_{6c}; however, we must reject H_{6d}.

Several main effects of trust and interactions of trust and differentiation strategy on resource investment achieve statistical significance only among either local or foreign parent firms; however, we find similar directional effects for the other parent as well. The statistical differences across the foreign and local parent firms appear to be superficial differences in effect sizes rather than statements of the nature or direction of the effects. Therefore, in larger samples, interorganizational trust, agency trust, and intraentity trust are likely to exhibit a consistent pattern of directional effects across both foreign and local parent firms.

Effects of Resource Investment and Utilization on Financial Performance

As we show in Table 4, in general, our findings support the RBV's theoretical perspective regarding the drivers of coentity performance. Financial performance is positively affected by the resource investments made by both parent firms (foreign: $\beta = .23$, $p < .01$; local: $\beta = .18$, $p < .05$) and the coentity's responsiveness ($\beta = .19$, $p < .05$); coentity coordination does not influence financial performance ($\beta = .10$, n.s.). Thus, we find support for H_7 and H_{8b}, but we must reject H_{8a} because the positive coefficient does not achieve statistical significance.[4]

We also find partial support for both H_{9a} and H_{9b}, in that resource investment's positive effect on financial performance is enhanced by both coordination (foreign: $\beta = .18$, $p < .05$; local: $\beta = .12$, n.s.) and responsiveness (local: $\beta = .20$, $p < .05$; foreign: $\beta = .11$, n.s.). Although two of these resource investment × resource utilization hypothesis tests do not achieve statistical significance, in all four cases, the coefficients are positive. In larger samples, these interactions likely would be detected across both foreign and local parent firms' investments. It is also possible that a more comprehensive assessment of resource investments that encompasses both the level and the quality of the parent firm's investments would generate more consistent results.

Discussion

Theoretical Contributions

Marketing literature has highlighted the critical role of trust (e.g., Atuahene-Gima and Li 2002; Morgan and Hunt 1994; Palmatier et al. 2006) and blames insufficient trust for the

---

[4] Of the control variables we examined, only dysfunctional competition affects the coentity's financial performance ($\beta = -.16$, $p < .05$).

TABLE 4
Determinants of Collaborative Entity (Coentity) Financial Performance

Variables	Hypotheses	Financial Performance (Standardized Regression Coefficient)
Control Variables		
Dysfunctional competition		−.16*
Environmental dynamism		.11
Cultural distance		.03
Years of operation		.09
Joint venture type (1 = vertical; 0 = horizontal)		.04
Industry (1 = business market; 0 = consumer market)		.10
Main Effects		
Resource investment (foreign parent)	H_7	.23**
Resource investment (local parent)	H_7	.18*
Coentity coordination	H_{8a}	.10
Coentity responsiveness	H_{8b}	.19*
Moderating Effects		
Resource investment (foreign parent) × coentity coordination	H_{9a}	.18*
Resource investment (local parent) × coentity coordination	H_{9a}	.12
Resource investment (foreign parent) × coentity responsiveness	H_{9b}	.11
Resource investment (local parent) × coentity responsiveness	H_{9b}	.20*
R^2		43
Adjusted R^2		.35
F-value		4.79**

*$p < .05$.
**$p < .01$.

poor results of collaborations, as well as for their failure to meet managers' expectations (Buchel 2003; Inkpen and Beamish 1997). Marketing and management researchers, including Doney and Cannon (1997) and Zaheer, McEvily, and Perrone (1998), note the importance of considering both interfirm (e.g., supplier and customer firms) and interpersonal (e.g., salesperson and buyer) factors to achieve full explications of the nature and effects of cross-firm ties. Therefore, we extend this line of inquiry by identifying three distinct levels at which trust operates within a coentity—between the collaborating firms, between a collaborating firm and its representatives, and among the members assigned by both firms to the coentity—and find that these three levels have distinct patterns of effects.

First, a collaborating firm's agency trust in its representatives increases the firm's resource investment in the coentity. Intraentity trust among the representatives assigned to the coentity affects resource utilization by both encouraging coordination and reducing responsiveness. The collaborating firms' interorganizational trust promotes their investment in the coentity and simultaneously enhances intraentity trust's positive effect on coordination and negative effect on responsiveness. These results imply that theoretical models of marketing collaborations and marketing relationships must carefully consider how forces operating at diverse levels can affect consequences. Particular attention should focus on the potential impact of factors at higher levels of aggregation; for example, interorganizational trust between the collaborating firms moderates the effects of intraentity trust on two constructs internal to the coentity, namely, coordination and responsiveness. Depending on the specific construct, relevant antecedents may exist at both the same and other levels. Therefore, limiting a theoretical focus to a single organizational level risks overestimating same-level effects and simultaneously missing important cross-level effects. A single-level focus may be problematic in any interorganizational marketing relationship, but it is especially troublesome in marketing collaborations that involve multiple firms, diverse constituents, partially conflicting agendas, mixed-motive situations, and complex divided loyalties.

A marketing collaboration involves a coentity when (1) two collaborating firms deliberately pursue a mutual collaboration to achieve one or more mutual goals; (2) each collaborating firm devotes and invests tangible and/or intangible resources in the collaboration, including assigning specific employees to the collaboration; and (3) the specific employees identified and assigned by each collaborating firm must engage with the other firm's representatives to plan, implement, and enact strategies to achieve the collaboration's mutual goals. The members of the collaboration constitute the coentity. Although we test our model using a sample of IJVs, we expect that the theoretical model and the underlying relationships among constructs generalize to strategic alliances, product development teams, project teams, and other marketing collaborations that involve the creation of a coentity.

For example, consider a product development team created by an automaker and a Tier 1 supplier. The collaborating firms must establish a sufficient level of interorganizational trust before the collaboration can begin or investments flow into it. If the firms have low levels of trust, they may incorporate other safeguards, such as nondisclosure agreements, and governance mechanisms, such as detailed reporting procedures, into the coentity's (product development team's) operations. Furthermore, to be successful, the product development team members need to share information, ideas, and insights and work together effectively. Because this sharing and coordination creates vulnerabilities that the partner might exploit, trust must be established within the team (i.e., intraentity trust). Over time, as the product development proceeds, the development team probably requires more resource input and more sensitive information from the collaborating firms. If the firms do not have sufficient agency trust in their respective project team members, they will not provide such information and resources. Therefore, deficient agency trust in the firm's own team members inhibits critical additional resource investments in the collaboration and results in suboptimal performance. Thus, as reflected in this example, all three types of trust are required for optimal coentity performance.

To examine trust at these three different levels, we focus on a specific trustor and specific trustee at each level. However, trust could also be examined in multiple ways that focus on different trustors and trustees at each level. At the interorganizational level, interorganizational trust could be operationalized and examined between the collaborating firms, or interpersonal trust could be operationalized and examined between the collaborating firms' chief executive officers. We could focus on interorganizational trust between a collaborating firm and the coentity, interpersonal trust between the collaborating firm's venture manager and the senior representative assigned to the coentity, or person-to-group trust between the collaborating firm's venture manager and the collective set of representatives assigned by the firm to the coentity. Intraentity trust assessments could include interpersonal measures of senior managers representing each parent or intergroup trust between the set of representatives assigned by each collaborating firm. Even more trustor–trustee configurations are possible; researchers' selection of a focal trustor–trustee relationship should be guided by theory and the specific set of antecedents, consequences, and other constructs under study. Our framework provides helpful guidance for identifying and examining relationships at multiple levels within complex collaborative entities.

Second, we demonstrate that building multilevel trust in isolation, without considering the relevant coentity's strategic and structural context, could be problematic. Although trust provides a motivation to engage in various positive behaviors, whether that motivation becomes manifest in actions depends on a host of factors, such as the coentity's formalized decision making and differentiation strategy. Specifically, formalized decision making not only increases the likelihood that representatives' intraentity trust motivates actual behaviors (i.e., parent firm's resource investments) but also enhances the negative inertial tendencies associated with intraentity trust. When the collaborative entity pursues a differentiation strategy, the insular tendencies generated by high intraentity trust are mitigated by the necessity of tuning in to the external environment. Our findings suggest that the more the coentity enacts a differentiation strategy, the more intraentity trust is manifest in coentity coordination. This recognition that a trusting party may be encouraged or prevented from acting on the motivation that arises from trust may help explain the inconsistency of trust's effects on performance outcomes, as revealed in a recent meta-analysis (Palmatier et al. 2006).

Third, in contrast to most previous studies, which suggest that trust building always leads to desirable outcomes (Dirks and Ferrin 2001), our research reveals that trust can be counterproductive in ways that extend beyond the obvious vulnerabilities discussed in previous research. The negative impact of intraentity trust on external responsiveness appears to be caused not by excessive vulnerability but rather by excessive closeness, insularity, and perhaps even a perception of invulnerability. In dynamic environments, the negative effects of intraentity trust—and how those effects may be exacerbated by trust at other levels, as well as by governance mechanisms—should be considered.

Managerial Implications

Our research offers several implications for managerial practice. Most extant research suggests that business executives should build trust with their partners to improve performance, but we consider this recommendation an oversimplification. When initiating a marketing collaboration, firms' decisions regarding coentity staffing, compensation decisions, and governance and management processes must balance the need for resource investment with the coentity's resource utilization requirements. If significant resource investments are required, maintaining a high level of agency trust is critical, and the selection and compensation of representatives could bolster agency trust by maximizing representatives' cultural similarity to the firm and tying part of their compensation to the firm's performance. Alternatively, if coordination is deemed to be the most critical, intraentity trust is paramount, and the selection and compensation of representatives should promote smooth and effective interactions with the partner's representatives, perhaps by linking compensation more closely to coentity performance.

The people who staff a coentity must balance their allegiance to their respective firms with their roles as colleagues of their counterparts. Because agency and intraentity trust have unique effects and operate through different mediators, collaborating members likely face situations in which their own firms' interests diverge from those of the coentity. In some cases, these representatives must choose between actions that build agency trust (and resource investment) and actions that increase intraentity trust, with its related improvements to the coentity's resource coordination. Selecting employees who have the capacity to deal with such conflicting loyalties is advisable.

Furthermore, managers can help ensure that trust-building efforts pay off by complementing relationship marketing activities with coentity structure and strategy. Specifically, if significant resource investments are needed, implementing formalized decision-making processes within the coentity will be particularly helpful because doing so can boost the effects of interorganizational trust on resource investment. In contrast, if high responsiveness is required, formalized decision making may worsen the negative effect of intraentity trust on resource responsiveness. The strategic context provides another factor worth considering because if a coentity pursues a differentiation strategy, the relative effects of multilevel trust on the coentity's resource investment and utilization are greater than if it adopts a low-cost strategy. Finally, building trust is not without drawbacks; intraentity trust may harm the coentity's external responsiveness, presumably because of the resultant excessive closeness and insularity. Thus, if it is important for a project team or strategic alliance to keep its edge, some of the employees assigned to that coentity should be rotated to deter excessive insularity.

Limitations and Future Research Directions

Although we develop our theoretical model to apply to all types of collaborative entities and believe that the underlying theoretical relationships are applicable to many types of marketing collaborations, our empirical sample is limited to IJVs. We have no empirical evidence that our specific results can be generalized to joint ventures in other contexts, strategic alliances, industry work groups, interfirm project teams, internal new product development task forces, or other collaborative entities. Only additional studies can address questions about generalizability.

We believe that the generalizability of some of our findings regarding formal decision-making procedures may be limited. The security provided by formal decision-making procedures appears to reduce concerns about exploitation and motivate greater investment by Chinese firms, but we do not find similar magnitude of effects for foreign collaborating firms. This moderating effect may be based in the underdevelopment of the institutional and legal infrastructure in China, which creates a more uncertain and riskier decision-making environment (Dahlstrom and Nygaard 1995). Thus, the interaction between agency trust and formalization of decision making and their effects on resource investment may be generalizable only to collaborating firms that face a similar lack of institutional protections. Further research should explore how the national or cultural background of the collaborating firms or the location of the coentity itself might affect resource investment, resource utilization, and, ultimately, performance.

Because we find that trust at each of the three levels has unique effects on performance, further exploration of their antecedents is worthwhile. Agency theory may shed light on the antecedents of agency trust, which could be affected by how representatives deal with the inherent mixed-motive situations of a coentity. To what extent do the representatives further the collaborating firm's objectives or act in accord with the coentity's unique interests, in the interests of their fellow representatives, or in their own self-interests? Channels and other interorganizational research might provide insights about trust building between parent firms, and management research could shed more light on the genesis of intraentity trust.

We focus on the effects of trust, but we also find that several control variables affect coentity performance. These findings point to avenues for further research. The significant effects of parent monitoring and interest alignment on resource investments and utilization suggest that these alternative governance mechanisms are worthy of attention, particularly in terms of whether trust's beneficial effects might be enhanced by some governance mechanisms, such as interest alignment, or whether trust can be undermined by other types of governance, such as monitoring programs. In addition, exploring whether the roles of trust at multiple levels differ in horizontal versus vertical collaborations might be fruitful; the potential for conflicts of interest multiply when collaborating firms are also competitors. Other factors, such as the coentity's life cycle, duration, leadership structure, and market environment, may leverage the impact of trust at various levels on resource investment and utilization.

Perhaps previous studies have focused too narrowly by investigating only the beneficial effects of trust. This implicit limited scope may have missed potential negative effects and limited the understanding of the full range of trust's effects. Therefore, additional research should investigate whether trust at various levels might have other negative ramifications.

APPENDIX

Constructs	Loading
Interorganizational Trust: Mutual Trust Between Collaborating Firms (Composite Reliability [CR] = .87)	
Both partners' parent companies trust each other.	.81
Both partners' parent companies are always frank and truthful in dealing with each other.	.83
Both parent companies believe that the other parent company would go out of its way to make sure the relationship is not damaged or harmed.	.71
Agency Trust: Collaborating Firm's Trust in Its Representatives Assigned to Coentity (CR = Foreign .75 and Local .74)	
(Please think of "we" and "us" as senior management representing your party in this joint venture and "our partner" as the senior management representing the other joint venture party identified at the beginning of the questionnaire)	
Our parent company trusts us in our ability to run this joint venture.	.75/.74
Our parent company trusts us in keeping their best interests in mind when running this joint venture.	.82/.81
Intraentity Trust: Mutual Trust Between Representatives of Different Firms Assigned to Coentity (CR = .85)	
(Please think of "we" and "us" as senior management representing your party in this joint venture and "our partner" as the senior management representing the other joint venture party identified at the beginning of the questionnaire)	
Our partner and we can rely on each other to do our job in the joint venture.	.76
In this relationship, our partner and we are both concerned about what happens to each other.	.67
When our partner and we share our problems with each other, both of us know that the other party will respond with understanding.	.74
Collaborating Firm's Resource Investment in Coentity (Formative)	
Please indicate the extent to which your parent organization has invested into the joint venture *after* the establishment of the joint venture.	
Product development, sales personnel and training, capital equipment, information systems, research and development, manufacturing and technology, marketing and distribution, and intellectual property (responses provided for all items)	N.A.
Coentity Coordination (CR = .93)	
Both partners play an active role in making important decisions such as product development, sales management, etc., for the joint venture	.73
Both partners consult with each other concerning important decisions regarding the joint venture.	.82
Both partners work effectively together to improve the operations of this joint venture.	.91
Both partners are always looking for synergistic ways to do business together.	.85
Coentity Responsiveness (CR = .93)	
Our joint venture can very quickly deploy our resources in places to face competitive actions.	.76
Our joint venture can quickly adjust our business process in response to environmental changes.	.81
Our joint venture can effectively sense the changes and trends in the marketplace, and act on them promptly.	.84
Coentity Formalization of Decision Making (CR = .89)	
Plans must be rigidly followed during implementation in our joint venture.	.74
There is a "standard operating procedure" for almost all major decisions.	.86
There are rules and procedures for most things.	.84
Our joint venture follows written procedures in most aspects of business in our joint venture.	.80
Coentity Differentiation Strategy (CR = .88)	
Our strategy can be described as:	
Maintaining a technological edge over competition in our market.	.76
Maintaining higher quality standards for our products.	.80
Maintaining a unique image for our products.	.71
Providing innovative products and/or service to our customers.	.82

APPENDIX
Continued

Constructs	Loading
Dysfunctional Competition (CR = .89)	
In the market of our principle industry, unlawful competitive practices such as illegal copying of new products are quite common.	.72
Our joint venture has experienced counterfeiting of our new products by other firms.	.73
There is no effective market competitive laws to protect our alliance's intellectual property.	.81
Our joint venture has experienced increased unfair competitive practices by other firms in the industry.	.82
Environmental Dynamism (CR = .88)	
In the market of our joint venture, customers' preferences change quickly over time.	.73
For our joint venture, market demand and consumer tastes have been unpredictable.	.80
Actions of local and foreign competitors of the joint venture have been highly unpredictable.	.77
The competition of our joint venture is changing very rapidly.	.75
Contract Specificity (CR = .84)	
Our relationship with our partner is governed primarily by written contracts.	.75
Everything is spelled out in detail in the contract of the joint venture.	.77
Over time, we have developed ways of doing things with our partner that never need to be expressed formally in the contract. (R).	.67
Parent Monitoring (CR = Foreign .82 and Local .80)	
Our parent company watches us closely to be sure that we keep their best interests in mind.	.77/.75
Our behaviors in the joint venture are monitored constantly by our parent company.	.80/.77
Our parent company requires us to report our handling of joint venture operation regularly.	.75/.72

Notes. All items were measured using seven-point scales anchored by 1 = "strongly disagree" and 7 = "strongly agree." R = reverse scored. N.A. = not available

REFERENCES

Aiken, Leona S. and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*. Thousand Oaks, CA: Sage Publications.

Atuahene-Gima, Kwaku and Haiyang Li (2002), "When Does Trust Matter? Antecedents and Contingent Effects of Supervisee Trust on Performance in Selling New Products in China and the United States," *Journal of Marketing*, 66 (July), 61–81.

Ayers, Doug, Robert Dahlstrom, and Steven J. Skinner (1997), "An Exploratory Investigation of Organizational Antecedents to New Product Success," *Journal of Marketing Research*, 34 (February), 107–116.

Bagozzi, Richard P., Youjae Yi, and Lynn W. Phillips (1991), "Assessing Construct Validity in Organizational Research," *Administrative Science Quarterly*, 36 (3), 421–58.

Bamford, James, David Ernst, and David G. Fubini (2004), "Launching a World-Class Joint Venture," *Harvard Business Review*, 82 (2), 90–100.

Barney, Jay B. (1991), "Firm Resources and Sustained Competitive Advantage," *Journal of Management*, 17 (1), 99–120.

Bonoma, Thomas V. and Wesley J. Johnston (1978), "The Social Psychology of Industrial Buying and Selling," *Industrial Marketing Management*, 7 (4), 213–24.

Buchel, Bettina (2003), "Managing Partner Relationships in Joint Ventures," *MIT Sloan Management Review*, 44 (4), 91–5.

Cannon, Joseph P. and Christian Homburg (2001), "Buyer–Supplier Relationships and Customer Firm Costs," *Journal of Marketing*, 65 (January), 29–43.

Child, John and Guido Mollering (2003), "Contextual Confidence and Active Trust Development in the Chinese Business Environment," *Organization Science*, 14 (1), 69–80.

Chiles, Todd H. and John F. McMackin (1996), "Integrating Variable Risk Preferences, Trust, and Transaction Cost Economics," *Academy of Management Review*, 21 (1), 73–99.

Currall, Steven C. and Andrew C. Inkpen (2002), "A Multilevel Approach to Trust in Joint Ventures," *Journal of International Business Studies*, 33 (3), 479–95.

Dahlstrom, Robert and Arne Nygaard (1995), "An Exploratory Investigation of Interpersonal Trust in New and Mature Market Economies," *Journal of Retailing*, 71 (4), 339–61.

Das, T.K. and Bing-Sheng Teng (1998), "Between Trust and Control: Developing Confidence in Partner Cooperation in Alliances," *Academy of Management Review*, 23 (3), 491–512.

Dirks, Kurt T. and Donald L. Ferrin (2001), "The Role of Trust in Organizational Settings," *Organization Science*, 12 (4), 450–67.

Doney, Patricia M. and Joseph P. Cannon (1997), "An Examination of the Nature of Trust in Buyer–Seller Relationships," *Journal of Marketing*, 61 (April), 35–51.

Dyer, Jeffrey H. and Harbir Singh (1998), "The Relational View: Cooperative Strategy and Sources of Interorganizational Competitive Advantage," *Academy of Management Review*, 23 (4), 660–79.

Eisenhardt, Kathleen M. and Jeffrey A. Martin (2000), "Dynamic Capabilities: What Are They?" *Strategic Management Journal*, 21 (10–11), 1105–1121.

Fornell, Claes and David F Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Gargiulo, Martin and Mario Benassi (2000), "Trapped in Your Own Net? Network Cohesion, Structural Holes, and the Adaptation of Social Capital," *Organization Science*, 11 (2), 183–96.

Gerbing, David W. and James C. Anderson (1988), "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment," *Journal of Marketing Research*, 25 (May), 186–92.

Ghosh, Mrinal and George John (1999), "Governance Value Analysis and Marketing Strategy," *Journal of Marketing*, 63 (Special Issue), 131–45.

Gilliland, David I. and Daniel C. Bello (2002), "Two Sides to Attitudinal Commitment: The Effect of Calculative and Loyalty Commitment on Enforcement Mechanisms in Distribution

Channels," *Journal of the Academy of Marketing Science*, 30 (1), 24–43.

Grabher, Gernot (1993), *The Embedded Firm: On the Socioeconomics of Industrial Networks*. New York: Routledge.

Grayson, Kent and Tim Ambler (1999), "The Dark Side of Long-Term Relationships in Marketing Services," *Journal of Marketing Research*, 36 (February), 132–41.

Hewett, Kelly and William O. Bearden (2001), "Dependence, Trust, and Relational Behavior on the Part of Foreign Subsidiary Marketing Operations: Implications for Managing Global Marketing Operations," *Journal of Marketing*, 65 (October), 51–66.

Inkpen, Andrew C. and Paul W. Beamish (1997), "Knowledge, Bargaining Power, and the Instability of International Joint Ventures," *Academy of Management Review*, 22 (1), 177–202

Jap, Sandy D. (1999), "'Pie-Expansion Efforts: Collaboration Processes in Buyer–Supplier Relationships," *Journal of Marketing Research*, 36 (November), 461–75.

——— and Shankar Ganesan (2000), "Control Mechanisms and the Relationship Life Cycle: Implications for Safeguarding Specific Investments and Developing Commitment," *Journal of Marketing Research*, 37 (May), 227–45.

Jaworski, Bernard J. and Ajay K. Kohli (1993), "Market Orientation: Antecedents and Consequences," *Journal of Marketing*, 57 (July), 53–70.

Jeffries, Frank L. and Richard Reed (2000), "Trust and Adaptation in Relational Contracting," *Academy of Management Review*, 25 (4), 873–82.

John, George (1984), "An Empirical Investigation of Some Antecedents of Opportunism in a Marketing Channel," *Journal of Marketing Research*, 21 (August), 278–89.

Johnson, Jean L., John B. Cullen, Tomoaki Sakano, and Hideyuki Takenouchi (1996), "Setting the Stage for Trust and Strategic Integration in Japanese-U.S. Cooperative Alliances," *Journal of International Business Studies*, 27 (5), 981–1004.

Kelly, Dawn and Terry L. Amburgey (1991), "Organizational Inertia and Momentum: A Dynamic Model of Strategic Change," *Academy of Management Journal*, 34 (3), 591–612.

Kogut, Bruce (1989), "The Stability of Joint Ventures: Reciprocity and Competitive Rivalry," *Journal of Industrial Economics*, 38 (December), 183–98.

——— and Harbir Singh (1988), "The Effect of National Culture on the Choice of Entry Mode," *Journal of International Business Studies*, 19 (3), 411–32.

Kumar, Sanjiv and Anju Seth (1998), "The Design of Coordination and Control Mechanisms for Managing Joint Venture-Parent Relationships," *Strategic Management Journal*, 19 (6), 579–99.

Lane, Peter J., Jane E. Salk, and Marjorie A. Lyles (2001), "Absorptive Capacity, Learning, and Performance in International Joint Ventures," *Strategic Management Journal*, 22 (12), 1139–61.

Li, Haiyang and Kwaku Atuahene-Gima (2001), "Product Innovation Strategy and the Performance of New Technology Ventures in China," *Academy of Management Journal*, 44 (6), 1123–34.

Loess, Kurt H. and Ugur Yavas (2003), "Human Resource Collaboration Issues in International Joint Ventures: A Study of U.S.-Japanese Auto Supply IJVs," *Management International Review*, 43 (3), 311–27.

Luo, Yadong (2000), *Multinational Corporations in China: Benefiting from Structural Transformation*. Copenhagen: Copenhagen Business School Press.

——— (2002), "Contract, Cooperation, and Performance in International Joint Ventures," *Strategic Management Journal*, 23 (10), 903–919.

——— (2005), "How Important Are Shared Perceptions of Procedural Justice in Cooperative Alliances," *Academy of Management Journal*, 48 (4), 695–709.

Makadok, Richard (2001), "Toward a Synthesis of the Resource-Based and Dynamic-Capability Views of Rent Creation," *Strategic Management Journal*, 22 (5), 387–401.

Maltz, Elliot and Ajay K. Kohli (1996), "Market Intelligence Dissemination Across Functional Boundaries," *Journal of Marketing Research*, 33 (February), 47–61.

McGraw, Kenneth and S.P. Wong (1996), "Forming Inferences About Some Intraclass Correlation Coefficient," *Psychological Methods*, 1 (1), 30–46.

Moorman, Christine, Gerald Zaltman, and Rohit Deshpandé (1992), "Relationships Between Providers and Users of Market Research: The Dynamics of Trust Within and Between Organizations," *Journal of Marketing Research*, 29 (August), 314–28.

Moran, Peter and Sumantra Ghoshal (1999), "Markets, Firms, and the Process of Economic Development," *Academy of Management Review*, 24 (3), 390–412.

Morgan, Neil A., Anna Kaleka, and Constantine S. Katsikeas (2004), "Antecedents of Export Venture Performance: A Theoretical Model and Empirical Assessment," *Journal of Marketing*, 68 (January), 90–108.

Morgan, Robert M. and Shelby D. Hunt (1994), "The Commitment–Trust Theory of Relationship Marketing," *Journal of Marketing*, 58 (July), 20–38.

Palmatier, Robert W., Rajiv P. Dant, and Dhruv Grewal (2007), "A Comparative Longitudinal Analysis of Theoretical Perspectives of Interorganizational Relationship Performance," *Journal of Marketing*, 71 (October), 172–94.

———, ———, ——— , and Kenneth R. Evans (2006), "Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis," *Journal of Marketing*, 70 (October), 156–53.

Porter, Michael E. (1985), *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: The Free Press.

Prahalad, C.K. and Gary Hamel (1990), "The Core Competency of the Corporation," *Harvard Business Review*, 68 (May–June), 79–91.

Reed, Richard and Robert J. DeFillippi (1990), "Causal Ambiguity, Barriers to Imitation, and Sustainable Competitive Advantage," *Academy of Management Review*, 15 (1), 88–102.

Rindfleisch, Aric (2000), "Organizational Trust and Interfirm Cooperation: An Examination of Horizontal Versus Vertical Alliances," *Marketing Letters*, 11 (1), 81–95.

——— and Christine Moorman (2001), "The Acquisition and Utilization of Information in New Product Alliances: A Strength-of-Ties Perspective," *Journal of Marketing*, 65 (April), 1–18.

Scheer, Lisa K. and Louis W. Stern (1992), "The Effect of Influence Type and Performance Outcomes on Attitude Toward the Influencer," *Journal of Marketing Research*, 29 (February), 128–42.

Schoorman, David, Roger Mayer, and James Davis (2007), "An Integrated Model of Organizational Trust: Past, Present, and Future," *Academy of Management Review*, 32 (2), 344–54.

Selnes, Fred and James Sallis (2003), "Promoting Relationship Learning," *Journal of Marketing*, 67 (July), 80–95.

Smith, J. Brock and Donald W. Barclay (1997), "The Effects of Organizational Differences and Trust on the Effectiveness of Selling Partner Relationships," *Journal of Marketing*, 61 (January), 3–21.

Snedecor, G.W. and W.G. Cochran (1980), *Statistical Methods*, 7th ed. Ames: Iowa State University Press.

Steenkamp, Jan-Benedict E.M. and Hans Baumgartner (1998), "Assessing Measurement Invariance in Cross-National Consumer Research," *Journal of Consumer Research*, 25 (June), 78–90.

Teece, David J., Gary Pisano, and Amy Shuen (1997), "Dynamic Capabilities and Strategic Management," *Strategic Management Journal*, 18 (7), 509–533.

Tsai, Wenpin and Sumantra Ghoshal (1998), "Social Capital and Value Creation: The Role of Intrafirm Networks," *Academy of Management Journal*, 41 (4), 464–76.

Tushman, Michael L. and Philip Anderson (1986), "Technological Discontinuities and Organizational Environments," *Administrative Science Quarterly*, 31 (3), 439–65.

Van Bruggen, Gerrit H., Gary L. Lilien, and Manish Kacker (2002), "Informants in Organizational Marketing Research: Why Use Multiple Informants and How to Aggregate Responses," *Journal of Marketing Research*, 39 (November), 469–78.

Wathne, Kenneth H. and Jan B. Heide (2000), "Opportunism in Interfirm Relationships: Forms, Outcomes, and Solutions," *Journal of Marketing*, 64 (October), 36–51.

——— and ——— (2004), "Relationship Governance in a Supply Chain Network," *Journal of Marketing*, 68 (January), 73–89.

Williamson, Oliver E. (1985), *The Economic Institutions of Capitalism*. New York: The Free Press.

Zaheer, Akbar, Bill McEvily, and Vincenzo Perrone (1998), "Does Trust Matter? Exploring the Effects of Interorganizational and Interpersonal Trust on Performance," *Organization Science*, 9 (2), 141–59.

Zaltman, Gerald and Thomas V. Bonoma (1977), "Organizational Buying Behavior: Hypotheses and Directions," *Industrial Marketing Management*, 6 (1), 53–60.

———, Robert Duncan, and Jonny Holbek (1973), *Innovations and Organizations*. New York: John Wiley & Sons.

Zellner, Arnold (1962), "An Efficient Method of Estimating Seemingly Unrelated Regression and Tests for Aggregation Bias," *Journal of the American Statistics Association*, 57 (298), 346–68.

Ji Hee Song & George M. Zinkhan

Determinants of Perceived Web Site Interactivity

Interactivity is a key feature of Web sites. This article identifies the determinants that enhance user perceptions of interactivity in a communication scenario in which consumers send instant messages to an e-store. Two conceptualizations of interactivity—telepresence theory and interactivity theory—predict that different antecedents (e.g., the number of clicks, response time, message type) are important. The results of Experiment 1 indicate that message type (i.e., how personal a particular message is) is the strongest predictor of interactivity perceptions. Furthermore, the findings suggest that the effects of message type on perceived interactivity and Web site effectiveness are greater when consumers are complaining than when they are inquiring about services. The results of Experiment 2 show that as the level of message personalization increases, interactivity perceptions and site effectiveness are enhanced (linear relationship). The authors discuss the implications of the findings for theory and practice and provide directions for measuring and manipulating interactivity in further research.

Keywords: interactivity, perceived interactivity, telepresence theory, Internet shopping, Web site design

Web sites are an important marketing tool for enhancing customer attraction, delivering service, facilitating transactions, and so forth. Prior research has investigated this medium as an interactive communication channel (Duncan and Moriarty 1998; Hoffman and Novak 1996), an e-shopping tool (Wolfinbarger and Gilly 2003), a business-to-business platform (Varadarajan and Yadav 2002), and a tool for customer relationship management (Payne and Frow 2005). Among the many important features associated with this medium, interactivity is unique compared with other mass media.

As technology evolves, there is a tendency for firms to harness the interactive capabilities of the Web. For example, Barwise and Farley (2005) find that the Web site is the most commonly used interactive marketing tool (e.g., as a public site, to sell products, for online chatting). Therefore, both marketers and Web site designers want to know how to build interactive Web sites. For example, why are some sites perceived as more interactive than others?

A possible answer derives from evolving technology. Recent research has suggested that increasing the quantity of some Web site features results in greater interactivity (Coyle and Thorson 2001; Macias 2003; Sicilia, Ruiz, and Munuera 2005). Such studies manipulate interactivity by varying the number of available features (e.g., presence of e-mail links, chat rooms, and search function; number of hyperlinks). However, we argue that the mere presence or absence of certain features matters only if these features affect how consumers navigate and use the site (Lee et al. 2004). For example, the presence of a chat room may not affect consumers' perceptions of interactivity unless they actually use the chat room. From a communication perspective, what are the best ways to enhance perceptions of a Web site's interactivity? Early work on interactivity has listed potential interactivity features, whereas more recent studies have focused on interactivity perceptions. However, little research has examined the link between interactivity features and perceptions. To investigate this issue, we use a specific scenario in which consumers visit a firm's Web site and initiate online chat to solve problems or gather information. By setting a limited context (i.e., consumers pursue a specific task), we create a situation in which consumers are motivated to use multiple features.

Steuer's (1992) telepresence theory and Rafaeli's (1988) interactivity theory provide a conceptual basis for identifying key features of interactivity perceptions. Although these two theories are frequently cited and used to conceptualize and operationalize interactivity, little empirical research has been conducted to test them. Telepresence theory posits that information is not merely transmitted from a sender to a receiver. Rather, according to this view, mediated environments are created and then experienced. Thus, both the properties of the medium (e.g., speed, range) and people's relationship to that medium contribute to interactivity perceptions. In contrast, Rafaeli views interactivity as a process of message exchange. According to interactivity theory, the quality of the communication (i.e., how reciprocal a particular exchange is) is the most important determinant of interactivity. Therefore, interactivity theory posits that message quality (i.e., the content of the response message) posi-

---

Ji Hee Song is Assistant Professor of Marketing, College of Business and Economics, University of Seoul, South Korea (e-mail: jiheesong@uos.ac.kr). (At the time of writing, she was an assistant professor, Department of Marketing and e-Business, College of Business and Economics, Towson University.) George M. Zinkhan is Coca-Cola Company Chair of Marketing, Department of Marketing, Terry College of Business, University of Georgia (e-mail: gzinkhan@terry.uga.edu). This research is based on the first author's dissertation research. The authors give special thanks to the anonymous *JM* reviewers and to Srinivas Reddy, Richard T. Watson, and Judy Harris for their helpful comments.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

tively influences users' perceptions of interactivity and site effectiveness.

Other factors that might affect interactivity perceptions include individual consumer characteristics and situations. Cognitive control theory (Averill 1973) argues that in stressful situations, people feel more in control when they are given more information. In the case of the current study, this theory predicts that people who send messages to a store about product delivery problems (i.e., a complaining situation) will have greater interactivity perceptions when they receive high-quality messages (e.g., personalized messages) than people who send product-inquiry messages.

Thus, the objectives of this article are threefold: (1) to identify the key Web site features that enhance interactivity perceptions; (2) to test two conceptualizations of interactivity, which suggest different antecedents of interactivity perception; and (3) to test the moderating effect of tasks in the formation of interactivity. We begin by reviewing the conceptualizations and operationalizations of interactivity. Then, we present a conceptual model and discuss the relationships among the key variables. Next, we describe the design of the experiment, which tests the hypotheses. Finally, we present our findings and their implications for the manipulation of interactivity and for managerial practice.

Interactivity Conceptualization

Before we identify the determinants of perceived interactivity, it is important to review what perceived interactivity is and how it is measured. Newhagen, Corders, and Levy (1995) were the first to suggest the concept of perceived interactivity. They viewed interactivity as a psychological variable in a content analysis of e-mail messages of viewers of the NBC news series *Almost 2001*. In their study, perceived interactivity is based on "efficacy," which is a two-dimensional construct: (1) viewers' psychological sense of efficacy and (2) viewers' sense of the media system's interactivity. That is, perceived interactivity is related to senders' beliefs that the NBC news series could process their message as useful input and then, in some way, act on it; thus, it corresponds to the idea of system efficacy. Subsequent studies have used these two dimensions to investigate perceived interactivity. For example, Wu (1999) applies these two dimensions to Web users, renaming them "internally based efficacy" and "externally based system efficacy." According to Wu, internally based efficacy can be translated into Web users' perceived control over where they are and where they are going. In contrast, externally based system efficacy can be rendered into users' sense of how responsive a Web site as a system is to their actions (Wu 1999, p. 255).

Prior research has identified an additional dimension of perceived interactivity called "direction of communication." These studies have applied either the traditional organizational communication model (Grunig and Grunig 1989) or the interpersonal communication model (Rafaeli 1988). Direction of communication is the extent to which users believe that the site facilitates two-way communication. In summary, perceived interactivity includes three dimensions (Liu 2003; McMillan and Hwang 2002): (1) control (internally based efficacy), (2) responsiveness (externally based system efficacy), and (3) communication.

To identify the key determinants of interactivity perception, we examine how interactivity concepts are operationalized and measured in previous empirical studies. Most studies have operationalized interactivity as the presence or absence of particular features. Thus, if one site has more options (e.g., choice of background colors, search option) than another site, the first site is judged to have a high degree of interactivity. Table 1 provides a list of extant studies that have identified or applied key interactive design features. As we mentioned previously, these features are often predicted by two key theories: telepresence theory and interactivity theory. For example, Coyle and Thorson (2001) manipulate interactivity on the basis of Steuer's (1992) definition: (1) the presence of a clickable image map and (2) the number of clickable areas. The features of feedback mechanisms in the communications dimension (e.g., e-mail links, chat rooms) are derived from interactivity theory. For example, Rafaeli (1988) suggests that the quality of message exchange is the key to interactivity. In adopting Rafaeli's theory, McMillan (2002) and Sicilia, Ruiz, and Munuera (2005) argue that the presence of two-way communication methods (e.g., e-mail links, toll-free telephone numbers) facilitates reciprocal relationships and subsequently affects interactivity perceptions. However, varying the quantity of these features may not be appropriate in manipulating interactivity. First, without valid manipulation checks, it is not possible to know which features are perceived as more interactive and which are perceived as less interactive (Liu and Shrum 2002). Second, interactivity depends on how users navigate the sites and which interactive features they actually use (Tremayne 2005). Therefore, adding interactive features (e.g., a chat room, e-mail links) to a site might not affect consumers' perceptions unless they actually use the features.

The problem of interactivity manipulation has been identified in several studies. For example, in her analysis of 108 health-related Web sites, McMillan (2002) finds no relationship between the interactive features and perceived interactivity. That is, respondents perceived some sites with more interactive features as not interactive, and vice versa. Furthermore, after comparing objective site characteristics and users' perceptions, Lee and colleagues (2004) find that though there was no significant difference among three Web sites on interactive features, participants rated one site significantly more interactive than the others. Thus, increasing the number or presence of site features does not necessarily stimulate an increase in interactivity.

Why is there a discrepancy between feature-based interactivity and perceived interactivity? What are the best ways to stimulate interactivity? The preceding review of the interactivity literature reveals some potential problems associated with manipulating and measuring interactivity. First, the potential site features that stimulate interactivity have not been empirically tested. In other words, the key interactivity theories (i.e., telepresence theory and interactivity theory) have often been used to define and measure interactivity, but they have not been validated or tested in an empirical setting. Second, Web site features and structures

TABLE 1
Summary of Key Interactive Features

Dimensions	Features	Examples of Site (System) Features	Authors
Communication	Feedback mechanisms	E-mail links, chat rooms, comment form, toll-free telephone number, question and answer, bulletin board, and frequently asked questions	Ghose and Dou (1998), Heeter (1998), Kiousis (2002), Macias (2003), Massey and Levy (1999), McMillan (2002), Rafaeli (1988), Sicilia, Ruiz, and Munuera (2005), Wu (2005)
	Message level	Percentage of overall messages based on prior message transmission	
	Transaction facilitation	Online order, order status tracking service	
	Information collection	Registration, cookies, and survey	Ha and James (1998), McMillan (2002)
Medium's structure	Options	Different languages, background color, searching option, news customization, and multimedia	Colye and Thorson (2001), Ghose and Dou (1998), Macias (2003), Massey and Levy (1999), McMillan (2000), Sicilia, Ruiz, and Munuera (2005), Tremayne and Dunwoody (2001), Wu (2005)
	Navigational choices	Search function, site map, clickable image map, and links	
Speed		Amount of time transmitting and taking messages (e.g., **response time, the number of clicks required to reach certain information**)	Kiousis (2002), Steuer (1992)
Others		Game, animation, and sensory complexity (amount of devices the system employs to activate the five senses)	Ha and James (1998), Kiousis (2002), Macias (2003)

Notes: In this study, we empirically study the features listed in bold.

can be viewed as antecedents of interactivity but not as interactivity itself, because most studies manipulated interactivity by varying the amount and nature of the structures (Tremayne 2005). Third, much of the interactive structure involves the presence or absence of simple features. In many cases, consumers were exposed to, but not specifically involved in using, these features. Fourth, relatively little research has included personal characteristics or situational variables that might make some features more salient than others.

In this article, we test three key features as potential antecedents of interactivity: the number of clicks, response time, and message type. We chose these features for three reasons. First, whereas feedback mechanisms and choice and navigation options have been discussed in prior research, relatively little attention has been paid to these features. Second, message type and speed are strongly predicted by both telepresence theory and interactivity theory. Third, message type and speed are particularly relevant to a computer-mediated communication scenario.

In contrast to previous approaches in which interactivity was manipulated by adding features, our study considers the quality of each feature. For example, we provide chat functions in all versions of the target sites. However, the quality of the chat function (e.g., speed, message quality) varies. In addition, participants are involved in using the relevant Web site features. For example, they must locate a chat button, initiate chatting, and wait for response messages.

Hypotheses Development

We apply telepresence theory and interactivity theory to identify three key determinants of interactivity perceptions. In addition, we apply cognitive control theory (Averill 1973) to specify hypotheses about the effect of different tasks on interactivity perceptions.

Telepresence Theory: Speed

As we mentioned previously, telepresence theory posits that a medium's structure influences users' sense of telepresence and medium usage (Steuer 1992). Steuer (1992, p. 76) defines telepresence as the extent to which a person feels engaged in a mediated environment (e.g., video game, television). For example, when people watch the news on television, they are experiencing the distant "real" world through a television monitor. Nonetheless, these viewers have the illusion of being there (i.e., a sense of telepresence).

Steuer (1992) theorizes that interactivity is a key aspect of telepresence, and he identifies important medium structures (capabilities) that contribute to interactivity and telepresence. Speed is one such medium structure;[1] it refers to the rate at which input can be assimilated into the mediated environment. Therefore, telepresence theory predicts that users will perceive more interactivity in a medium when that medium manages their requests/inputs quickly (see Figure 1). In our empirical setting, participants are asked to send instant messages to a store. Therefore, we define speed in terms of how quickly consumers can send instant messages and then how quickly they receive a response. On the basis of this definition, we operationalize speed in terms of two features: (1) the number of clicks required to reach the store's "live chat" button (i.e., how easily consumers find the button) and (2) response time (i.e., how quickly consumers receive a reply). Participants send inputs (clicks) to locate a store's "live chat" button. Therefore, they perceive more interactivity when they click fewer tiers to connect to information because a fewer number of clicks are associated with how quickly their inputs are assimilated into the medium (i.e., Web site). After participants send instant messages to an e-store, they must wait for the store's response. Thus, consumers' perceptions of interactivity and site effectiveness are enhanced as the store's response time to their instant messages decreases (H_1 and H_2). Finally, we expect that perceived interactivity mediates the relationship between a site's quick management of users' requests and site effectiveness. Telepresence theory posits that a medium's structure affects consumers' interactivity perceptions and behaviors (site effectiveness). Moreover, prior empirical research has shown that there is a positive relationship between interactivity perceptions and attitude or behavioral intentions (McMillan 2002; McMillan and Hwang 2002; Wu 1999). Therefore, we propose mediating effects in H_3. Again, perceived interactivity includes three dimensions: (1) perceived communication, (2) perceived control, and (3) perceived responsiveness. We measure site effectiveness in terms of six variables: attitude, site quality, loyalty, satisfaction, word of mouth (WOM), and repurchase behavior. We operationalize "quick management" as the number of clicks it takes to locate a chat button and the response time from the store.

H_1: A Web site's quick management of users' requests has a positive effect on (a) perceived communication, (b) perceived control, and (c) perceived responsiveness.

H_2: A Web site's quick management of users' requests has a positive effect on their (a) satisfaction with the site, (b) loyalty, (c) attitude toward the site, (d) perceptions of site quality, (e) WOM behavior, and (f) repurchase behavior.

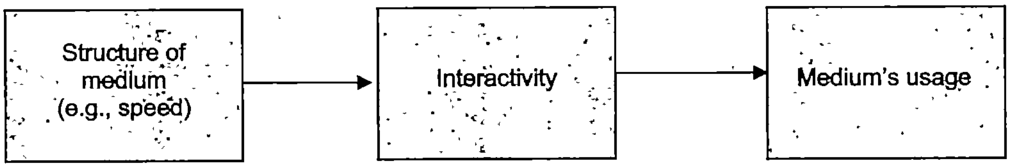
H_3: Perceived interactivity mediates the positive influence of a Web site's quick management of users' request on their (a) satisfaction with the site, (b) loyalty, (c) attitude toward the site, and (d) perceptions of site quality.

Interactivity Theory: Message Type

We hypothesize that message type is related to consumers' perceptions of interactivity and site effectiveness. In our empirical setting, participants receive response messages from an e-store. We apply interactivity theory to predict the relationship. Rafaeli (1988) distinguishes among three communication processes: (1) two-way communication, (2) reaction, and (3) interaction. A two-way communication is established as soon as a message flows bilaterally (i.e., the presence of contact information, such as toll-free telephone numbers and e-mail addresses). For example, as Figure 2 shows, as long as a person is able to send a message (M2) to an object (site), a two-way communication is facilitated. Reactive communication requires that subsequent messages refer to earlier ones (e.g., a response to someone's e-mail message). For example, a person sends a message (M1) to an object (site) and receives a response (M2) from the site. Interaction differs from reaction through the incorporation of previous references to the content, nature, or form of the message (e.g., the content of a replying e-mail is related to the previous messages). Thus, when a site sends a message that is related to a consumer's prior message, perceptions of interactivity and site effectiveness are enhanced. In Figure 2, an interactive message is based on the way subsequent messages (M2) are related to previous messages (M1). Similar to H_3, we expect that perceived interactivity mediates the effect of a site's transmission of related messages on site effectiveness. Thus:

H_4: A Web site's transmission of messages associated with previous exchanges (i.e., personalized message) has a

FIGURE 1
Telepresence Theory

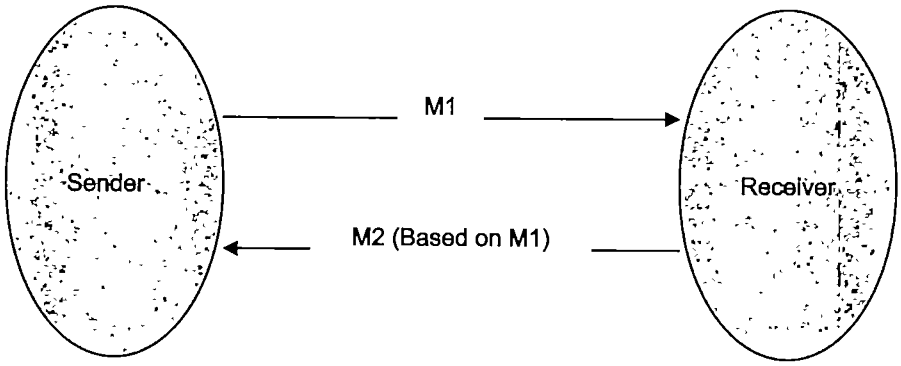


Notes: Speed is the rate at which input can be assimilated into the mediated environment

---

[1] Steuer (1992) identifies two additional features—mapping and range—that may stimulate interactivity. However, manipulating and testing all three features in one empirical setting is complicated, especially under the condition in which participants are required to use all features. In addition, speed is particularly appropriate in our communication setting. In contrast, mapping is more applicable for virtual reality or video games (Kiousis 2002), and range is often manipulated in terms of the navigational choices/options given to users. However, McMillan (2002) finds that the relationships between some range features (e.g., search function, choice) and interactivity perception are insignificant.

FIGURE 2
Interactivity Theory



Notes: M1 denotes message, and M2 denotes responding message.

positive effect on (a) perceived communication, (b) perceived control, and (c) perceived responsiveness.

H_5: A Web site's transmission of messages associated with previous exchanges has a positive effect on consumers' (a) satisfaction with the site, (b) loyalty, (c) attitude toward the site, (d) perceptions of site quality, (e) WOM behavior, and (f) repurchase behavior.

H_6: Perceived interactivity mediates the positive influence of a Web site's transmission of messages associated with previous exchanges on consumers' (a) satisfaction with the site, (b) loyalty, (c) attitude toward the site, and (d) perceptions of site quality.

Tasks

Are message effects (i.e., personalized versus standardized messages) similar when consumers have different tasks? In our experimental setting, we consider two tasks: search (prepurchase) and complaining (postpurchase). Under the prepurchase condition, participants send instant messages and request information about sales tax charges related to an e-store product (i.e., a T-shirt bearing a university logo). Under the postpurchase condition, participants submit a complaint about problems related to product delivery. In this case, we develop our hypotheses on the basis of cognitive control theory (Averill 1973). Cognitive control has two mechanisms: information gain and reappraisal of a stressful situation. Both mechanisms involve the cognitive efforts that people expend to cope with stressful situations. In the case of information gain, a person perceives the situation as more predictable; in the case of a reappraisal, the event is considered through a cognitive reinterpretation "to conform to the needs and desires of the individual" (Averill 1973, p. 293). That is, people impose meaning to the event. For example, under a stressful situation, consumers perceive the situation as more acceptable and controllable (reinterpretation) if they are given more information (information gain). Therefore, participants who complain about receiving the wrong product will perceive the site as more interactive and effective (reinterpretation) if they receive personalized messages (i.e., information gain or a message based on a previous message) rather than standardized messages. On the basis of this logic, we hypothesize the following:

H_{7a}: Perceived interactivity is greater for the complaining task (the most stressful situation) than for the search task (the least stressful situation) when the site transmits messages that are associated with previous messages.

H_{7b}: Perceived interactivity is lower for the complaining task (the most stressful situation) than for the search task (the least stressful situation) when the site transmits messages that are not associated with previous messages.

H_{8a}: Site effectiveness is greater for the complaining task (the most stressful situation) than for the search task (the least stressful situation) when the site transmits messages that are associated with previous messages.

H_{8b}: Site effectiveness is lower for the complaining task (the most stressful situation) than for the search task (the least stressful situation) when the site transmits messages that are not associated with previous messages.

Experiment 1

Stimulus and Participants

We tested the hypotheses in an experimental setting. The experiment was a 2 (number of clicks) × 2 (response time) × 2 (message type) × 2 (tasks) between-subjects factorial design. Thus, there were 16 versions of a Web site. The stimulus was a fictitious, Web-based store (Webstore.com) that sells university-related souvenirs and paraphernalia (e.g., T-shirts, mugs). We conducted the experiment in a computer lab. Participants were recruited from undergraduate business classes and were scheduled to participate in a 30-minute laboratory session. Participants received class credit as a motivation or incentive for their participation. In addition, we entered all the students into five raffle drawings of $30 each. A total of 341 participants took part in the experiment, and we randomly assigned them to the 16 treatment levels. All participants were required to have at least some online experience; however, online shopping experience was not a requirement for participation. We eliminated 5 participants for failure to demonstrate serious responses. Of the remaining 336 participants, 46.4% were female, 83.6% were white, 29.9% had a household income of $80,000 or lower, and their mean age was 20.6 years (range of 18–32 years).

Procedure

Participants were introduced to the purpose of the study as follows: "Webstore.com is a start-up, Web-based retailer that sells university-related souvenirs. The company's target market is young people, especially college students like yourself. Webstore.com plans to launch its first Web site in the fall of 2006. Prior to launch, the company wants to test its Web site with potential customers." Then, participants read one of the following scenarios:

> Search scenario (prepurchase): Your mother's birthday is coming soon, and you'd like to buy her something. You've decided to buy a T-shirt (with a university logo, shown below) in Webstore.com. Before you make a purchase decision, you want to know whether the store charges any tax on your order. So, you would like to contact the store and ask this. To obtain a quick response, you want to engage in live chatting with a store representative.[2]
>
> Complaining scenario (postpurchase): About one week ago, you purchased a birthday gift for your mother. Specifically, you purchased a T-shirt bearing a university logo. Yesterday, you received the item and discovered that the store sent the wrong T-shirt. Now, you want to know how to receive the right item as soon as possible and how to return the wrong item. So, you would like to contact the store. To obtain a quick response, you want to engage in live chatting with a store representative.

After reading one of these scenarios, participants were asked both to visit the e-store and to initiate a live chat with the e-store (i.e., send questions through instant messenger). Participants then waited until they received a response. We collected perceptual dependent measures, such as perceptions of interactivity and site effectiveness (i.e., satisfaction, loyalty, attitude, and site quality), and obtained two behavioral effectiveness measures (i.e., WOM behavior and repurchase behavior). First, we asked participants to log on to their preferred e-mail accounts and to send e-mail messages to their best friends about their shopping experiences. We then collected the e-mails in an alias e-mail account that we created for the study. Second, we gave participants four similar sites (three "real" e-stores and our target e-store) and asked them to choose one site to purchase a gift after they looked around the four sites.[3] Finally, participants answered some demographic questions.

Independent Variables Manipulation

Number of clicks. We manipulated the number of clicks in terms of how many clicks were required to locate the "live chat" button of the target store. Under the condition with the fewest clicks, participants needed to click one time to locate the button. Under the condition with many clicks, participants needed to click at least six times to locate the button.

We conducted a manipulation check before the final experiment. Participants answered manipulation check questions for perceived ease of use (Davis, Bagozzi, and Warshaw 1989). We used seven-point Likert-type scales to indicate agreement or disagreement with the following two statements: "Finding the 'live chat' button is easy," and "Navigating the site is easy for me."

Response time. We manipulated response time in terms of how quickly participants received a reply from the store about their inquiries or complaints. Under the slow condition, participants received a reply after 90 seconds. Under the fast condition, participants received a reply after 15 seconds.

To validate the stimulate level of response time, we measured perceived quickness (Gorn et al. 2004). We used seven-point Likert-type scales to assess users' perceptions of "slow/fast," "not speedy/speedy," and "not quick/quick."

Message type. Message type is the extent to which the reply is based on prior messages. Under the baseline condition, participants received a standardized message. For the highly personalized message, the store responded to a participant's request appropriately. For example, in the case of a complaining situation (i.e., a wrong product delivery), participants received a message that specifically addressed their requests. Therefore, we manipulated message type as follows: (1) Participants received a standardized message, or (2) participants received a personalized message. Under the standardized message condition in all scenarios, participants received the following e-mail: "Thanks for contacting us. Support is currently unavailable. Please provide your name and e-mail address, and we will get back to you shortly." Under the personalized message condition in Scenario 1, participants received the following e-mail: "We are required by law to charge sales or use tax in any state in which it has a physical presence. The actual tax rate applied to your purchase is based on the destination of the order." Finally, under the personalized message condition in Scenario 2, participants received the following e-mail: "First, we apologize for the inconvenience. You will get the right item no later than tomorrow evening. For your convenience, we have provided a preaddressed UPS return label. Please drop off the return item at the UPS store nearest you. Of course, you do not need to pay anything. Do you have any more questions?"

Participants answered the manipulation check questions for perceived personalization (Wolfinbarger and Gilly 2003). We used a seven-point Likert-type scale to indicate agreement or disagreement with the following two statements: "The message gives me personal attention," and "The message understands my specific needs."

Dependent Variables

We used existing measures of perceived interactivity (Liu 2003; McMillan and Hwang 2002; Wu 1999). Specifically, we measured perceived interactivity on three dimensions: (1) communication (six items), (2) control (nine items), and

---

[2] This phrase might introduce bias because it primes participants to think that a quick response is desirable. Thus, we conducted a test to examine the effect (to obtain a quick response) of the phrase on participants' expectations of service speed. The result shows that the phrase had no significant effect on participants' expectations of service.

[3] All four sites sold school-related souvenirs; however, some variance among the four sites might exist in terms of product variety, overall site design, and prices. For example, our target site did not have transaction functions (e.g., shopping basket, payment). Participants knew that the target store was testing its Web site with potential customers, so the lack of a transaction function did not affect their decisions.

(3) responsiveness (six items) (see Appendix A). The coefficient alphas computed for the three dimensions ($\alpha = .934$, .901, and .922, respectively) indicate high internal consistency. We measured site effectiveness with four perceptual measures (i.e., satisfaction, loyalty, attitude toward the Web site, and site quality) and two behavioral measures (i.e., repurchase behavior and WOM). For repurchase behavior, participants were given several alternative stores, including the target e-store, and were asked which store they would visit in the future. For WOM communications, participants were asked to send e-mail messages to their friends about their overall shopping experiences at the e-store. The coefficient alphas computed for the four perceptual measures indicate strong internal consistency: (1) satisfaction (three items; $\alpha = .928$), (2) loyalty (five items; $\alpha = .956$), (3) attitude (three items; $\alpha = .950$), and (4) quality (two items; $\alpha = .942$). Appendix A provides a complete description of the measurement approaches.

Pretest

We conducted the first pretest to check the manipulations of the four variables: (1) situation (normal versus stressful), (2) number of clicks (one versus six), (3) response time (15 seconds versus 60 seconds), and (4) message type (standardized versus personalized). Sixteen students participated in the pretest. The results show significant differences between a normal situation and a stressful situation, between one click and six clicks, and between a standardized message and a personalized message. However, there is no significant difference between the 15-seconds delay and the 60-seconds delay. On the basis of this result, we manipulated the response time as 15 seconds versus 90 seconds.

Before we ran the pilot study, 10 graduate students checked the overall quality of the stimulus and questionnaire, and we made some changes in response to their suggestions. The pilot study validated the treatment levels and overall experimental process. We exposed 40 students (who were randomly assigned to eight cells) to the total experimental experience under the same set of conditions as those of the main experiment. The results from the pilot study provide support for the internal validity of the experiment. Participants showed significant differences in their perceptions of situation (t = –4.899, d.f. = 38, $p < .05$), clicks (t = 7.102, 4.897; d.f. = 38, $p < .05$), response time (t = 4.872, 3.899, 4.154; d.f. = 38, $p < .05$), and message types (t = –5.816, –7.777; d.f. = 38, $p < .05$). Regarding the response time manipulation, there was a significant difference in the perceptions of speed between 15 seconds and 90 seconds. The pilot study findings indicate that no other changes were needed for the main experiment.

Findings

Number of clicks. Before examining the effect of the number of clicks on the individual dependent measures, we conducted a multivariate test to assess the effects collectively. Table 2 provides summary outputs from the multivariate analysis of variance (MANOVA). All four multivariate differences measures (i.e., Pillai's trace, Hotelling's trace, Wilks' lambda, and Roy's largest root) are significant ($p < .05$); that is, all the dependent variables (i.e., interactivity perception, satisfaction, loyalty, attitude, and quality) vary across the different number of clicks. Given the significance of the multivariate test, we examined the result of a univariate test (within MANOVA) to determine whether all the dependent variables were significantly different or whether the results were derived from differences of only several dependent variables. The main effects of the number of clicks on two dimensions of perceived interactivity (i.e., control and responsiveness) were significant (see Table 2). Participants in the one-click condition perceived the site as being more controllable (M = 5.51) than participants in the six-click condition (M = 4.41). Similarly, the number of clicks was a significant antecedent of responsiveness. In contrast, the effect of the number of clicks on communication perception was not significant. The findings provide notable insights into interactivity. When participants were able to locate certain information with minimum effort (i.e., a small number of clicks), they believed that they had control over their actions/requests (perceived control) and, at the same time, that the site was responsive to their requests and quick to respond. However, participants did not neces-

TABLE 2
MANOVA Result: Main Effects and Interaction Effects

Variable	d.f.	Number of Clicks: F-Value (Effect Size)	Response Time: F-Value (Effect Size)	Message Type: F-Value (Effect Size)	Message Type × Tasks (Interaction Effect): F-Value (Effect Size)
Interactivity 1 (communication)	1	2.444 (.008)	7.831* (.024)	168.993* (.346)	7.604* (.023)
Interactivity 2 (control)	1	96.776* (.232)	.014 (.000)	34.980* (.099)	3.460 (.011)
Interactivity 3 (responsiveness)	1	5.406* (.017)	85.734* (.211)	77.866* (.196)	2.752 (.009)
Satisfaction	1	22.356* (.065)	11.540* (.035)	221.354* (.409)	11.167* (.034)
Loyalty	1	34.513* (.097)	2.085 (.006)	90.741* (.221)	9.223* (.028)
Attitude	1	32.510* (.092)	2.681 (.008)	77.975* (.196)	5.472* (.017)
Quality	1	34.296* (.097)	5.700* (.018)	127.849* (.285)	9.110* (.028)

*$p < .05$.

sarily believe that the site facilitated two-way communication. Therefore, H_1 is partially supported.

There were significant effects of the number of clicks on satisfaction ($F(1, 320) = 22.356$, $p < .05$; mean square [MS] = 38.453), loyalty ($F(1, 320) = 34.513$, $p < .05$; MS = 58.674), attitude ($F(1, 320) = 32.510$, $p < .05$; MS = 50.298), and quality ($F(1, 320) = 34.296$, $p < .05$; MS = 50.298). The effect sizes of the site effectiveness measures were between 6.5% and 9.7%.

To investigate the possible mediating effect of interactivity perception on site effectiveness, we conducted a multivariate analysis of covariance (MANCOVA) with the number of clicks as an independent factor, site effectiveness measures (satisfaction, loyalty, attitude, and quality) as the dependent variables, and three interactivity perceptions as covariates. The results indicate that all the interactivity perceptions (covariates) were statistically significant ($p < .05$). However, the previously significant effect of the number of clicks on satisfaction disappeared when we included the interactivity perceptions as covariates ($F(1, 317) = 2.752$, $p > .05$; MS = 2.010). Moreover, all effect sizes for site effectiveness measures decreased when the interactivity perceptions were covariates. The findings provide evidence for the mediating role of interactivity perception. We compute the magnitude of the mediation as the percentage of reduction in the mean square of the effect produced by the interactivity perception (covariate) (Gorn et al. 2004). Interactivity perceptions mediated 84.2%–94.7% of the mean square effect for the number of clicks on site effectiveness measures: satisfaction = 94.7%, loyalty = 87.7%, attitude = 89.3%, and quality = 84.2%. That is, the number of clicks negatively affected site effectiveness, and this effect was mediated by the greater interactivity perceptions induced by a small number of clicks than a large number of clicks. Therefore, H_3 is supported.

Response time. The fourth column of Table 2 provides the result of a main effect of response time (the second manipulation of speed). The multivariate test reveals that all the dependent variables were significantly different across response time. The main effects of response time on two dimensions of perceived interactivity (i.e., communication and responsiveness) were significant. Participants in the fast-response condition perceived the site as more responsive (M = 4.612) than participants in the slow-response condition (M = 3.227). The effect of response time on the communication perceptions showed the same pattern. In contrast, the effect of response time on the control dimension was not significant ($F(1, 320) = .101$, n.s.). The findings indicate that when a site sends a replying message quickly, participants perceive the site as responsive and communicative. However, participants do not necessarily believe that they have control over their experiences. This result validates the prior finding that speed is a partial predictor of perceived interactivity.

There were significant effects of message response time on satisfaction ($F(1, 320) = 11.540$, $p < .05$; MS = 19.850) and quality ($F(1, 320) = 5.700$, $p < .05$; MS = 8.360). However, message response time was not a significant antecedent of loyalty and attitude. A possible reason for this could be the presence of the live chat function on the site. For example, participants were impressed that the site provided a real-time chatting function. In other words, the presence of a chatting function provided a positive image of this particular Web site, even though the store sent a slow reply (90 seconds). Therefore, the store's slow response had a negative effect on participants' specific experience with the store (in terms of both satisfaction and quality of experience). In contrast, a slow response did not affect participants' overall attitude or commitment to the site (loyalty). Thus, the findings partially support H_2.

To investigate the possible mediating effect of interactivity perceptions on site effectiveness, we conducted a MANCOVA, with response time as an independent factor, site effectiveness measures (satisfaction, loyalty, attitude, and quality) as the dependent variables, and three interactivity perceptions as covariates. The results indicate that all the interactivity perceptions (covariates) were statistically significant ($p < .05$). However, the effect sizes for satisfaction and quality measures decreased when we included interactivity perceptions as covariates. Interactivity perceptions mediated more than 50% of the mean square effect for the number of clicks on satisfaction (80.3%) and quality (58.4%). That is, response time negatively affected site effectiveness, and this effect was mediated by the greater interactivity perceptions induced by a fast response. Therefore, H_3 is supported.

Message type. The fifth column of Table 2 provides the results of the main effect of message type. The multivariate test reveals that all the dependent variables were significantly different across message type. The main effects of message type on three dimensions of perceived interactivity (i.e., communication, control, and responsiveness) were significant. Participants who received a personalized message perceived the site as more communicative, controllable, and responsive (M = 5.417, 5.278, and 4.558, respectively) than participants who received a standardized message (M = 3.547, 4.649, and 3.281, respectively). Moreover, the effect sizes were large (between 10% and 35%), which indicates that message type was a strong predictor of all the interactivity variables. The findings indicate that a site's transmission of related messages increases not only the perceptions of communication but also the perceptions of control and responsiveness. In other words, when a site transmits messages that are associated with previous exchanges (i.e., personalized message), consumers believe that they have more control over the site and, at the same time, that the site is responsive to their requests. There were significant effects of message type on satisfaction ($F(1, 320) = 221.354$, $p < .05$; MS = 380.730), loyalty ($F(1, 320) = 90.741$, $p < .05$; MS = 156.893), attitude ($F(1, 320) = 77.975$, $p < .05$; MS = 120.640), and quality ($F(1, 320) = 127.849$, $p < .05$; MS = 187.503). Therefore, H_4 and H_5 are supported.

We then conducted a MANCOVA with message type as an independent factor, site effectiveness measures (satisfaction, loyalty, attitude, and quality) as the dependent variables, and three interactivity perceptions as covariates. The results indicate that all the interactivity perceptions (covariates) were statistically significant ($p < .05$). In addition, the previously significant effects of message type on loyalty (F(1, 317) = 2.548, $p > .05$; MS = 2.415) and attitude (F(1, 317) = 1.221, $p > .05$; MS = 1.109) disappeared when we included the interactivity perceptions as covariates. Moreover, all effect sizes of the site effectiveness measure decreased when the interactivity perceptions were covariates. The findings provide evidence for the mediating role of interactivity perception. We compute the magnitude of the mediation as the percentage of reduction in the mean square of the effect produced by the interactivity perceptions (covariates). Interactivity perceptions mediated between 89.7% and 99.1% of the mean square effect for message type on the site effectiveness measures: satisfaction = 89.7%, loyalty = 98.4%, attitude = 99.1%, and quality = 94.7%. Again, this suggests that interactivity perceptions serve as a mediator between message type and site effectiveness. Therefore, H_6 is supported.

Interaction effect. The sixth column of Table 2 provides the results associated with the interaction effect. The multivariate test reveals that all the dependent variables were significantly different across the message types, by situations. The interaction effects of message type and situation on the communication dimension were significant. Participants who received a personalized message perceived the site as more communicative when they were under a stressful situation (i.e., wrong product delivery) (M = 5.543) than when they were under a less stressful situation (i.e., price inquiry) (M = 5.291). Similarly, participants who received a standardized message perceived the site as less communicative when they were under a stressful situation (i.e., wrong product delivery) (M = 3.293) than when they were under a less stressful situation (i.e., price inquiry) (M = 3.800). However, the interaction effect on the control and responsiveness dimension was not significant (F(1, 320) = 3.460, $p > .05$; F(1, 320) = 2.752, n.s., respectively). Therefore, H_{7a} and H_{7b} are partially supported. There were significant interaction effects on satisfaction (F(1, 320) = 11.167, $p < .05$; MS = 19.207), loyalty (F(1, 320) = 9.223, $p < .05$; MS = 15.947), attitude (F(1, 320) = 5.472, $p < .05$; MS = 8.466), and quality (F(1, 320) = 9.110, $p < .05$; MS = 13.360). Therefore, H_{8a} and H_{8b} are supported.

We then conducted a MANCOVA with message type and situation as independent factors, site effectiveness measures (satisfaction, loyalty, attitude, and quality) as the dependent variables, and three interactivity perceptions as covariates. The results indicate that all the interactivity perceptions (covariates) were statistically significant ($p < .05$). However, the previously significant result of the multivariate test was no longer significant. All four multivariate differences measures were not significant (F(1, 253) = 1.662, $p > .05$), indicating that interactivity perceptions mediated the interaction effect of message type and situation on site effectiveness.

WOM communication. We conducted content analysis for the 325 e-mail messages that participants sent to the site. Two coders were involved in the coding process. Each coder independently classified each e-mail message onto the coding sheets. We calculated interjudge reliability on the basis of PRAM (Program for Reliability Assessment with Multiple Coders). For the 18 items coded, intercoder reliability (raw percentage agreement) was 94.04%, with individual reliabilities ranging from 80.4% to 99.4%, which indicated excellent agreement (Neuendorf 2002). For more details on the 18 items and reliabilities, see Appendix B. The two coders resolved all disagreements; thus, 100% agreement was achieved.

We conducted cross-tabulation analysis to estimate the effect of site features on WOM behaviors. Participants in the one-click condition were more likely to exhibit positive WOM behavior than those in the six-click condition (χ^2 = 18.728, d.f. = 2, $p < .05$). In addition, participants who received personalized message were more likely to have a positive evaluation of the store than those who received standardized messages (χ^2 = 55.650, d.f. = 2, $p < .05$). However, response time did not affect participants' overall evaluations of the store. This finding suggests that the number of clicks and the types of messages received from the store are major concerns among e-shoppers. However, response time was not a strong predictor of overall store evaluations. This is similar to the prior result; that is, response time does not affect loyalty or attitude measures (e.g., overall feelings, commitment to the site).

To examine the effect of interactivity perceptions on WOM behavior, we conducted a MANOVA. In addition, to analyze pairwise differences between groups of sites, we conducted a post hoc Scheffé test. The participants who recommended Webstore.com to their friends had significantly higher communication perceptions (F(1, 322) = 80.871, $p < .05$), control perceptions (F(1, 322) = 39.391, $p < .05$), and responsiveness perceptions (F(1, 322) = 68.599, $p < .05$) than other shoppers. These results provide convergent evidence for the positive relationship between interactivity perceptions and site effectiveness (i.e., WOM behavior).

Repeat purchase behavior. In the final task, participants were asked to choose one site among four similar Web stores (including the target site) to purchase self-gifts. Of the 336 participants, 39 (11.6%) chose the target site (Webstore.com), 58 (17.3%) chose Site A, 129 (38.4%) chose Site B, and 110 (32.7%) chose Site C.

First, we conducted cross-tabulation to test the effect of site features (i.e., number of clicks, response time, and message type) on repeat purchase behavior. The message type strongly affected participants' repurchase behavior (χ^2 = 4.902, d.f. = 2, $p < .05$). However, the number of clicks and response time did not influence repeat purchase behavior. These findings suggest that message relatedness (i.e., how much the shop's message reflected participants' concerns) is an important feature affecting patronage behavior.

Second, to examine the effect of interactivity perceptions on repurchase behavior, we conducted a MANOVA. In addition, to analyze pairwise differences between the groups, we conducted a post hoc Scheffé test. The participants who chose the target site over the other three competing stores had significantly greater communication perceptions (F(1, 334) = 13.641, $p < .05$), control perceptions (F(1, 334) = 8.506, $p < .05$), and responsiveness perceptions (F(1, 334) = 23.072, $p < .05$) than the other participants. These results provide further evidence for the positive relationship between interactivity perceptions and site effectiveness (i.e., repurchase behavior).

Experiment 2

We designed Experiment 2 to explore further the message-type effect (the strongest predictor of interactivity perception). Experiment 1 manipulated message type with two extremes (a personalized reply versus a standardized reply). This is unrealistic in that everyday communication provides considerable variation in terms of the level of message personalization. In Experiment 2, we added a medium-personalized message and ran a between-subjects design with three experimental conditions: standardized message, medium-personalized message, and highly personalized message. The medium-personalized message did not specifically address participants' concerns (i.e., it was less personalized than the personalized message); however, it included some cues or signals that created a sense that participants' questions would be attended to and handled by the store in near future (i.e., it was more personalized than the standardized message).

Experiment 2 extends the generalizability of the message-type effect by using different stimuli. The personalized message manipulated in Experiment 1 could have been more personalized. For example, participants in the highly personalized condition were provided with a sales tax policy rather than an actual tax rate (based on their zip codes). In Experiment 2, we used a slightly different scenario and stimuli in which participants asked about the store's student discount rate. In addition, participants in the personalized-message condition were given the actual student discount rate. Thus, in Experiment 2, we anticipated that an increase in the level of personalization of the transmitting messages would enhance perceptions of interactivity and site effectiveness (linear trend).

Method

A total of 121 undergraduate students participated in Experiment 2 in exchange for extra credit. We randomly assigned participants to one of the three experimental conditions: (1) standardized message, (2) medium-personalized message, and (3) highly personalized message. The procedure and stimuli were similar to those of Experiment 1. The primary changes involved the use of a different scenario and a different message-type manipulation. The participants were given the following scenario:

> Your mother's birthday is coming soon, and you'd like to buy her something. You've decided to buy a T-shirt (with a school logo) from Webstore.com. Before you make a purchase decision, you want to know whether the store gives any discounts to students like you. So, you would like to contact the store to ask about a discount for students. Now, you want to engage in live chatting with a store representative.

We manipulated message type as follows: Under the standardized message condition, participants received the following e-mail: "Thanks for contacting us. Support is currently unavailable. Please provide your name and e-mail address, and we will get back to you shortly." Under the medium-personalized-message condition, participants received the following e-mail: "Hi, my name is Anna. Thanks for contacting us. I am currently taking care of another customer. Could you wait for a second? I will get back to you shortly." Finally, under the highly-personalized-message condition, participants received the following e-mail: "Hi, my name is Anna. Thanks for contacting us. We offer an instant 10% discount on all your purchases when you enter your student identification number. Do you have any other questions?"

The dependent measures were similar to those in Experiment 1: three measures of interactivity and four measures of site effectiveness (WOM and repurchase behavior were not measured). We performed manipulation checks. Participants rank-ordered "personal attention" as follows: (1) highly personalized message (M = 4.89), (2) medium-personalized message (M = 2.56), and (3) standardized message ($M = 1.6$; $F(1, 25) = 12.971$, $p < .05$). The extent to which the messages reflected participants specific needs was rank ordered as follows: (1) highly personalized message (M = 5.4), (2) medium-personalized message (M = 1.8), and (3) standardized message ($M = 1.4$; $F(1, 25) = 40.781$, $p < .05$).

Results

We used a MANOVA to examine the effect of message type on perceived interactivity and site effectiveness. The multivariate test indicated that all the dependent variables were significantly different across message types. The effect of message type on perceived communication was rank ordered as follows: the standardized message (M = 3.923), the medium-personalized message (M = 5.053), and the highly personalized message ($M = 5.68$; $F(2, 118) = 18.662$, $p < .05$). The results show a similar pattern for perceived control ($F(2, 118) = 3.527$, $p < .05$) and perceived responsiveness ($F(2, 118) = 9.829$, $p < .05$). There were significant linear effects of message type on satisfaction ($F(2, 118) = 14.571$; $p < .05$), loyalty ($F(2, 118) = 13.996$, $p < .05$), attitude ($F(2, 118) = 12.113$, $p < .05$), and quality ($F(2, 118) = 11.548$, $p < .05$).

To test the mediating role of interactivity perception, we conducted a MANCOVA, with the three interactivity perceptions as covariates. The results show that the previously significant effect of message type on site effectiveness measures disappeared, confirming the mediating role of interactivity perception.

In general, the results of Experiment 2 are consistent with those of Experiment 1. In particular, we find a linear relationship between the personalization level of the message and interactivity perceptions. In other words, interactivity perceptions are enhanced as the level of message personalization increases.

Discussion

In this article, we identify the key features that enhance consumer perceptions of interactivity. Specifically, we tested this relationship under a situation in which users chatted with an e-shopping site. Two conceptualizations of interactivity suggest different antecedents of interactivity perceptions. Telepresence theory posits that a property of a medium (i.e., speed) is an important determinant of interactivity perception. In contrast, interactivity theory posits that message quality is an antecedent of a Web site's interactivity. In addition, in applying cognitive control theory, we expected that the different tasks (e.g., search, complaints) would moderate interactivity formation. We report some evidence that interactivity theory is more successful in explaining interactivity perceptions (in both Experiments 1 and 2) than telepresence theory. The results suggest that the content of the message is the strongest predictor of perceptions of interactivity and site effectiveness. In addition, our findings support the interaction hypothesis; that is, the message-quality effects on both interactivity perceptions and site effectiveness vary under different user tasks (in support of cognitive control theory).

The findings help us form some tentative answers to the question, What are some feasible ways to enhance interactivity? As we mentioned previously, there are two main views of interactivity: telepresence theory and interactivity theory. Interactivity theory predicts the effect of message type (the strongest antecedent) on interactivity perceptions (see H_4). Our results provide strong support for this conceptualization of interactivity. Specifically, we find that message exchange is a stronger determinant of interactivity than are specific features of the medium.

The strong support for the interactivity theory could result from our specific empirical setting (i.e., communication scenario). Therefore, our results do not necessarily suggest that telepresence theory should be rejected. The two conceptualizations are not necessarily mutually exclusive. For example, Web designers who want to determine how users interact with the contents of a Web page may need to address the site's interface issues from the telepresence view. In contrast, interactivity theory is useful for describing how to incorporate computer-mediated communication channels (e.g., e-mail) as customer-service touchpoints. For example, almost half of firms' communications with their customers occur through online channels, such as e-mail and instant messaging (Anton and Phelps 2002). Thus, marketers who want to manage and maintain interactive communications with customers through the Internet need to consider the importance of enhancing message quality, as interactivity theory suggests.

The significant interaction effects in Experiment 1 support cognitive control theory (Averill 1973). According to this theory, under complex conditions, personalized messages (i.e., more information) strengthen cognitive reappraisal effects (i.e., consumers reinterpret the situation) and thus increase the positive evaluation of a site. That is, when consumers receive personalized messages, they report greater interactivity perceptions under a complaining task (a wrong product delivery) than under a search task (asking about product price).

We found that response time is a significant predictor of satisfaction and quality perceptions but not of loyalty, attitude, WOM behavior, and repurchase behavior. A possible reason for this result is that satisfaction and quality are experience-specific measures (i.e., measures about specific Web site experiences). Conversely, loyalty, attitude, WOM behavior, and repurchase behavior are overall measures (i.e., measures of overall feelings about a site). Because consumers valued the presence of a chatting function, they showed positive attitudes toward a site, regardless of the speed of the response. Thus, additional research that examines the effect of time on interactivity perceptions in different communication media (e.g., e-mail) is needed.

Our approach contributes to the literature on measuring/manipulating interactivity. In prior studies, interactivity has been manipulated as the presence or absence of certain e-scape features (e.g., presence of a site map). Moreover, participants have not necessarily been involved in using the particular features. In our setting, we manipulate interactivity in a way that allows participants to get involved in using the target features (e.g., navigating the site to find the chat function, initiating the chatting function, receiving response messages). On the basis of our empirical findings, we suggest the following two recommendations for manipulating and measuring Web site interactivity in further research: First, researchers should ensure that participants are involved in using the particular features of interest. Interactivity resides in the consumers' eyes, not in the system itself. Second, simply adding features does not guarantee a high level of interactivity. For example, one site might have five clickable links on its home page, and another site might have only two links. In line with prior studies, the former is considered interactive. However, if a consumer is looking for a firm's address, it is possible that none of the five links in the first site will help him or her find the information. In contrast, one of the two links on the second site might guide the user to the firm's address. Thus, the user will perceive the second site as more interactive. Therefore, when manipulating interactivity, researchers should consider both the quantity and the quality of the features. Similarly, researchers should be aware that some features are only partial predictors of interactivity. That is, some features might stimulate only one or two dimensions of interactivity.

Managerial Implications

The findings highlight the importance of managing and defining an e-servicescape (e.g., the physical surroundings of a Web site, the Web design). Key site-design features provide cues that stimulate positive perceptions of the site and encourage visitors to judge the effectiveness of the site. Our results suggest some ways that organizations can increase interactivity perceptions on their Web sites. In particular, our results indicate that some key design features (i.e., number of clicks, response time, and message type) are antecedents of consumers' interactivity perceptions. However, we found that merely including a certain feature (e.g., a chatting function) does not necessarily enhance interactivity. In our empirical setting, we added live chat functions to 16 versions of a site. Users' interactivity perceptions varied on each site, depending on the content or quality users derived from the chatting function. In other words, the quality of the chatting message (e.g., speed, the content of message) was a more important predictor of interactivity perceptions than the mere presence of a chat function. Thus, e-stores should be careful when adding interactive features to their Web sites because the mere presence of certain features does not necessarily enhance interactivity. One way to ensure the desired effect of interactive features would be to pretest key features in a usability study (Maity and Peters 2005). That is, e-stores could employ usability studies to learn more about consumers' overall experiences and, in turn, strive to improve Web site design by examining consumers' interactions with particular Web site features.

Of the three features we studied in this article, message type had the greatest effect on interactivity and effectiveness (Table 2 shows the effect sizes for the three interactivity measures and four site effectiveness measures). Specifically, increasing the level of message relatedness enhanced consumers' self-efficacy (perceived control) and system efficacy (perceived responsiveness of the site). Thus, when customers communicated with firms, the personalization of the seller's message was a crucial aspect of interactivity. In contrast, the number of clicks and response time were relatively less important features. This finding indicates that e-stores should carefully plan and monitor their customer communications. An impediment to computer-mediated communication is the lack of personalization and social appeal. Experiment 2 provides managers with some promising tactics for enhancing personalization. For example, the use of social presence cues (e.g., including greeting messages or a customer's name) in a message is one strategy that firms could use to increase the level of message relatedness.

Our study provides marketers with some ways to manage interactivity. Perceived interactivity includes three dimensions: (1) perceived communication, (2) perceived control, and (3) perceived responsiveness. The empirical results suggest some ways to enhance each dimension. For example, to increase perceived control, site designers should provide quick navigation (fewer clicks to locate specific information) and transmit messages associated with previous exchanges (e.g., providing relevant answers to users' questions, delivering personalized pages based on prior transactions). Similarly, fast response and transmission of related messages might be crucial for enhancing perceived communication. All three features (i.e., the number of clicks, response time, and message type) are strong predictors of perceived responsiveness, which indicates that managing perceived responsiveness is relatively easier than managing the other two dimensions of interactivity.

In our study, interactivity perceptions mediated more than 80% of the mean square effect for design features on site effectiveness. This finding supports our argument that creating and managing interactivity in computer-mediated communication is crucial for marketing managers.

Our findings suggest that customers' tasks and situations affect the relationship between message type and interactivity perceptions. Consider the situation in which consumers navigate an e-shopping site to find answers to problems such as a wrong delivery or returns and exchanges. An e-store should communicate with these customers with special care, using hotlines (i.e., a direct number without waiting) or fast, real-time chatting services. Customers with general inquiries about products or services can be treated in cost-effective ways, such as through e-mail communication, automated response systems, frequently asked questions, and a customer service page. In some cases, this approach might create challenges for marketers. As we show, depending on the situation (e.g., complaining versus searching), consumers approach an e-store with different motivations to process the interactive features that they find there. In some banking sites, for example, online chatting windows pop up when consumers are navigating the site. Such a feature might be helpful for consumers who need immediate assistance (e.g., opening a checking account). In contrast, for consumers who are navigating the site to obtain financial information, the interactive power of real-time chatting might be of little use or even counterproductive. In other words, such searchers might be irritated by the chatting pop-up and subsequently leave the site. Therefore, it is important for marketers to understand the relevant situational factors that affect consumers' acceptance and adoption of specific Web site features.

Limitations and Further Research

Many site-design features may affect interactivity perceptions (see Table 1). In this study, we test only three features. Further research could test other features, such as personalized functions, navigational choices, and other methods that facilitate transactions. Herein, we considered a service situation in which customers interacted with a store by sending e-messages. Other studies should be conducted in a transaction setting, such as the search for or purchase of products. For example, some e-stores provide "my virtual model" in which consumers can "try on" the products of interest (e.g., clothing, furniture, shoes) to determine what they would look like. Many e-stores are using three-dimensional pictures to help consumers make better decisions. It would be worthwhile to test the role of these sensory features in shaping interactivity perceptions.

The use of a social presence cue (i.e., a sense of each party's presence) might stimulate interactivity perceptions (Short, Williams, and Christie 1976). For example, in a classroom setting, teachers know that they are interacting with students when they send social presence cues (e.g., eye contact, facial expression). Similarly, when users communicate through a Web site and the site facilitates social presence cues (i.e., creating a sense of communicating with a "real" person), they will perceive the medium as interactive. Further research might explore the effect of related cues (e.g., accent, use of a familiar name). For example, Indian call-center operators are trained to disguise their local accents, replacing them with American, Canadian, or British patterns of speech (Friedman 2005, p. 26).

Most studies on service waiting time have been conducted in traditional service contexts (e.g., a face-to-face context). As a result, waiting times are measured in minutes or hours rather than in seconds. It is not clear how online customers evaluate waiting time in comparison with in-person waiting times. Therefore, it would be worthwhile to examine the relationship between site or store features and interactivity perceptions under different communication channels, including face-to-face, in-store, chatting, e-mail, telephone, and Web pages. In addition, customers' tolerance for waiting might be another important factor that affects interactivity perceptions. For example, customers who are willing to wait two minutes for a response during a chat communication may show different interactivity perceptions than consumers who are willing to wait for only 20 seconds. It would also be worthwhile to examine consumers' perceived waiting duration. For example, consider a situation in which a customer is waiting for a service on the Internet (e.g., sending an instant message). A user who is involved in other activities (e.g., surfing the Internet, checking e-mail messages) during the waiting period might perceive a shorter waiting duration than a user who is waiting for the service without the distraction of the other activities.

Users' experiences with shopping sites influence interactivity perceptions. Thus, longitudinal studies would contribute valuable insights. For example, it would be interesting to know how consumers' interactivity perceptions evolve as they accumulate more experiences with a particular site or with a particular category of sites. Herein, we attempted to incorporate objective measures of site effectiveness (e.g., shoppers sent e-mail messages to friends). Further research should examine other objective measures (e.g., the number of hits, the number of clicks, time spent on the site).

From an organizational perspective, it is important to consider both benefits and costs when modifying a Web site's design. For example, the addition of a chatting function provides benefits to consumers. However, the associated costs (e.g., generating prompt responses) are a barrier to organizational adoption. Further research should more directly study the effect of design features on firm performance (e.g., return on investment).

In summary, using two key conceptualizations of interactivity, we identified some important features affecting interactivity perceptions. Our findings provide evidence that executional factors (e.g., the number of clicks, message type) affect perceived interactivity. In turn, we found that interactivity is an important predictor of site effectiveness.

APPENDIX A
Measures of Dependent Variables

Interactivity Perceptions (Seven-Point Scale) (Liu 2003; McMillan and Hwang 2002; Wu 1999)

Communication

1. This Web site facilitates two-way communication.
2. The Web site gives me the opportunity to talk back.
3. The Web site facilitates concurrent communication.
4. The Web site enables conversation.
5. The Web site does not encourage visitors to talk back.R
6. The site is effective in gathering visitors' feedback.

Control

1. While I was on the site, I was always aware where I was.
2. While I was on the site, I always knew where I was going.
3. While I was on the site, I was always able to go where I though I was going.
4. I was delighted to be able to choose which link and when to click.
5. I feel that I have a great deal of control over my visiting experience at this site.
6. The Web site is not manageable.R
7. While I was on the site, I could choose freely what I wanted to see.
8. While surfing the site, I had absolutely no control over what I could do on the site.R
9. While surfing the site, my actions decided the kind of experiences I got.

Responsiveness

1. The Web site processed my input very quickly.
2. Getting information from the Web site is very fast.
3. I was able to obtain the information I want without any delay.
4. When I clicked on the links, I felt I was getting instantaneous information.
5. The Web site was very slow in responding to my request.R
6. The Web site answers my question immediately.

Attitude Toward the Site (Seven-Point Scale) (Coyle and Thorson 2001)

1. Good/bad
2. Favorable/unfavorable
3. Like/dislike

Satisfaction (Seven-Point Scale) (Fornell et al. 1996)

1. I am satisfied with the experience.
2. This online shopping experience is exactly what I needed.
3. This online experience hasn't worked out as well as I thought it would.R

Overall Site Quality (Seven-Point Scale) (Wolfinbarger and Gilly 2003)

1. The overall quality of my purchase experience with this Web site is ... (1 = "very poor," and 7 = "excellent").
2. My overall feelings toward this Web site are ... (1 = "very dissatisfied," and 7 = "very satisfied").

Loyalty Intentions (Seven-Point Scale) (Zeithaml, Berry, and Parasuraman 1996)

1. I encourage friends and relatives to do business with the Web site.
2. I say positive things about the Web site to other people.
3. I will do more business with the Web site in the next few years.
4. I would recommend the Web site to someone who seeks my advice.
5. I consider this Web site to be my first choice to buy the kind of product I most recently purchased online.

Repurchase (Pan 2003)

Participants choose one Web-based store out of four possible stores.

WOM (Song, Zinkhan, and Pan 2005)

Participants write e-mail messages about shopping experiences to their best friends.

Notes: R = reverse coding

APPENDIX B
Intercoder Reliability

	Items	Percentage Agreement		Items	Percentage Agreement
Positive comments	Easy to find chat function	80.4	Negative comments	Easy to find chat function	98.4
	Message speed	92.5		Message speed	93.8
	Message content	93.5		Message content	93.8
	Navigation	94.1		Navigation	97.5
	Price	98.1		Price	99.4
	Selection	89.8		Selection	98.8
	Overall feeling	97.5		Overall feeling	95.3
	Design	95.7		Design	94.4
	Recommend	87.6		Not recommend	92.2
			Overall		94.04

REFERENCES

Anton, Jon and Dru Phelps (2002), “How to Benchmark Your Call Center,” (accessed October 20, 2006), [available at http://www.benchmarkportal.com/newsite/index.tml].

Averill, James (1973), “Personal Control over Aversive Stimuli and Its Relationship to Stress,” *Psychological Bulletin*, 80 (4), 286–303.

Barwise, Patrick and John U. Farley (2005), “The State of Interactive Marketing in Seven Countries: Interactive Marketing Comes of Age,” *Journal of Interactive Marketing*, 19 (3), 67–80.

Coyle, James R. and Esther Thorson (2001), “The Effects of Progressive Levels of Interactivity and Vividness in Web Marketing Sites,” *Journal of Advertising*, 30 (3), 65–77.

Davis, Fred D., Richard P. Bagozzi, and Paul R. Warshaw (1989), “User Acceptance of Computer Technology: A Comparison of Two Theoretical Models,” *Management Science*, 35 (8), 982–1003.

Duncan, Tom and Sandra E. Moriarty (1998), “A Communication-Based Marketing Model for Managing Relationships,” *Journal of Marketing*, 62 (April), 1–13.

Fornell, Claes, Machael D. Johnson, Eugene W. Anderson, Jaesung Cha, and Barbara E. Bryant (1996), “The American Customer Satisfaction Index: Nature, Purpose, and Findings,” *Journal of Marketing*, 60 (October), 7–18.

Friedman, Thomas L. (2005), *The World Is Flat: A Brief History of the Twenty-First Century*. New York: Farrar, Straus and Giroux.

Ghose, Sanjoy and Wenyu Dou (1998), “Interactive Functions and Their Impacts on the Appeal of Internet Presence Sites,” *Journal of Advertising Research*, 38 (2), 29–43.

Gorn, Gerald J., Amitava Chattopadhyay, Jaideep Sengupta, and Shashank Tripathi (2004), “Waiting for the Web: How Screen Color Affects Time Perception,” *Journal of Marketing Research*, 41 (May), 215–25.

Grunig, James E. and Larissa A. Grunig (1989), “Toward a Theory of Public Relations Behavior of Organizations: Review of a Program of Research,” in *Public Relations Research Annual*, Vol. 1, J.E. Grunig and L.A. Grunig, eds Hillsdale, NJ: Lawrence Erlbaum Associates, 27–63.

Ha, Louisa and Lincoln James (1998), “Interactivity Reexamined: A Baseline Analysis of Early Business Web Sites,” *Journal of Broadcasting and Electronic Media*, 42 (4), 457–74.

Heeter, Carrie (1989), “Implications of New Interactive Technologies for Conceptualizing Communication,” in *Media Use in the Information Age: Emerging Patterns of Adoption and Computer Use*, J.L. Salvaggio and J. Bryant, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 217–35.

Hoffman, Donna L. and Thomas P. Novak (1996), “Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations,” *Journal of Marketing*, 60 (July), 50–68.

Kiousis, Spiro (2002), “Interactivity: A Concept Explication,” *New Media & Society*, 4 (3), 355–83.

Lee, Se-Jin, Wei-Na Lee, Hyojin Kim, and Patricia Stout (2004), “A Comparison of Objective Characteristics and User Perception of Web Sites,” *Journal of Interactive Advertising*, 4 (2), [available at http://www. jiad.org].

Liu, Yuping (2003), “Developing a Scale to Measure the Interactivity of Websites,” *Journal of Advertising Research*, 43 (3), 207–16.

——— and L.J. Shrum (2002), “What Is Interactivity and Is It Always Such a Good Thing? Implications of Definition, Person, and Situation for the Influence of Interactivity on Advertising Effectiveness,” *Journal of Advertising*, 31 (4), 53–64.

Macias, Wendy (2003), “A Beginning Look at the Effects of Interactivity, Product Involvement and Web Experience on Comprehension: Brand Web Sites as Interactive Advertising,” *Journal of Current Issues and Research in Advertising*, 25 (2), 31–44.

Maity, Moutusi and Cara Lee Okleshen Peters (2005), “A Primer for the Use of Internet Marketing Research Tools: The Value of Usability Studies,” *Interactive Marketing*, 6 (3), 232–45.

Massey, Brian L. and Mark R. Levy (1999), “Interactivity, Online Journalism and English-Language Web Newspapers in Asia,” *Journalism and Mass Communication Quarterly*, 76 (1), 138–51.

McMillan, Sally J. (2000), “Interactivity Is in the Eye of the Beholder: Function, Perception, Involvement, and Attitude Toward the Web Site,” in *Proceedings of the 2000 Conference of the American Academy of Advertising*, Mary Alice Shaver, ed. East Lansing: Michigan State University, 71–78.

——— (2002), “A Four-Part Model of Cyber-Interactivity: Some Cyber-Places Are More Interactive Than Others,” *New Media and Society*, 4 (2), 271–91.

——— and Jang-Sun Hwang (2002), “Measures of Perceived Interactivity: An Exploration of the Role of Direction of Communication, User Control, and Time in Shaping Perceptions of Interactivity,” *Journal of Advertising*, 31 (3), 29–41.

Neuendorf, Kimberly A. (2002), *The Content Analysis Guidebook*. Thousand Oaks, CA: Sage Publications.

Newhagen, John E., John W. Corders, and Mark R. Levy (1995), “Nightly@nbc.com: Audience Scope and the Perception of Interactivity in Viewer Mail on the Internet,” *Journal of Communication*, 45 (3), 164–75.

Pan, Yue (2003), "Online and Offline Patronage Behavior in a Retail Setting: Three Studies," doctoral dissertation, Department of Marketing, University of Georgia.

Payne, Adrian and Pennie Frow (2005), "A Strategic Framework for Customer Relationship Management," *Journal of Marketing*, 69 (October), 167–76.

Rafaeli, Sheizaf (1988), "Interactivity: From New Media to Communication," in *Advancing Communication Science: Merging Mass and Interpersonal Processes*, R.P. Hawkins, J.M. Wiemann, and S. Pingree, eds. Newbury Park, CA: Sage Publications, 110–34.

Short, John A., Ederyn Williams, and Bruce Christie (1976), *The Social Psychology of Telecommunications*. New York: John Wiley & Sons.

Sicilia, Maria, Salvador Ruiz, and Jose L. Munuera (2005), "Effects of Interactivity in a Web Site," *Journal of Advertising*, 34 (3), 31–46.

Song, Ji Hee, George M. Zinkhan, and Yue Pan (2005), "Young Shoppers' Experience with E-Shopping," *International Journal of Internet Marketing and Advertising*, 2 (3), 220–37.

Steuer, Jonathan (1992), "Defining Virtual Reality: Dimensions Determining Telepresence," *Journal of Communication*, 42 (4), 73–93.

Tremayne, M. (2005), "Lessons Learned from Experiments with Interactivity on the Web," *Journal of Interactive Advertising*, 5 (2), [available at http://www.jiad.org].

——— and S. Dunwoody (2001), "Interactivity, Information Processing, and Learning on the World Wide Web," *Science Communication*, 23 (2), 111–34.

Varadarajan, P. Rajan and Manjit Yadav (2002) "Marketing Strategy and the Internet: An Organizing Framework," *Journal of the Academy of Marketing Science*, 30 (4), 296–312.

Wolfinbarger, M. and M.C. Gilly (2003), "eTailQ: Dimensionalizing, Measuring and Predicting eTail Quality," *Journal of Retailing*, 79 (3), 183–98.

Wu, Guohua (1999), "Perceived Interactivity and Attitude Toward Websites," in *Proceedings of the American Academy of Advertising*, M.S. Roberts, ed. Gainseville: University of Florida, 254–62.

——— (2005), "The Mediating Role of Perceived Interactivity in the Effect of Actual Interactivity on Attitude Toward the Website," *Journal of Interactive Advertising*, 5 (2), [available at http://www.jiad.org].

Zeithaml, Valarie A., Leonard L. Berry, and A. Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60 (April), 31–46.

Alina B. Sorescu & Jelena Spanjol

Innovation's Effect on Firm Value and Risk: Insights from Consumer Packaged Goods

What is the relationship between innovation and firm value? Does the type of innovation make a difference? To answer these questions, the authors examine how breakthrough and incremental innovations affect three different facets of firm performance: normal profits, economic rents, and total firm risk. They argue that each of these metrics is of independent interest to shareholders and managers and that examining one without the others results in an incomplete picture of the true financial value of innovation. Using data on more than 20,000 new products from consumer packaged goods industries, the authors find that breakthrough innovation is associated with increases in both normal profits and economic rents and that, on average, each breakthrough innovation in the sample is associated with an increase in firm value of $4.2 million. Breakthrough innovation is also associated with increases in the risk of the innovating firm, but this higher risk is offset by above-normal stock returns. In contrast, incremental innovation is associated with increases in normal profits only and has no impact on economic rents or firm risk.

Keywords: innovation, shareholder value, firm risk, marketing strategy, marketing–finance interface

Many executives hold an unwavering belief in innovation as a strategic imperative, counting on innovation to spur growth and yield positive financial returns. However, profitable innovation remains an elusive goal. Chief executive officers (CEOs), such as Sun Microsystems' Jonathan Schwartz (2006), recognize innovation as "the key to survival," but they are still wondering how to get innovation to pay off. A Boston Consulting Group (2005) study concurs; whereas 74% of 940 executives surveyed expected to spend more on innovation in 2005 than in previous years, more than half the respondents were dissatisfied with the returns on their investments.

The academic literature is similarly muddled. Some authors suggest that simply innovating more will lead to better performance (Bayus, Erickson, and Jacobson 2003; Pauwels et al. 2004). Others argue that incremental innovations, which make up 90% of new product introductions, have little or no impact on firm value (Christensen 1997; Foster and Kaplan 2001). Indeed, executives who have chosen to focus on quantity find themselves lamenting their decision: "There's actually an innovation glut. The real shortage is profits" (Schrage 2000, p. 225). Moreover, some authors decry the riskiness of innovation and even caution managers that "going after breakthrough innovations may be glamorous, but mounting evidence suggests that's the last growth strategy you should try" (Treacy 2004, p. 29).

We attempt to clarify the apparent contradiction between studies that highlight the positive impact of innovation on firm value and those that argue the contrary. Following previous research, we differentiate between incremental and breakthrough innovations (e.g., Dewar and Dutton 1986; Srinivasan et al. 2006). We also recognize that firm performance has multiple facets and that the effect of innovation on these distinct facets can differ markedly. Specifically, we examine the impact of breakthrough and incremental innovations on firms' (1) normal profits, (2) economic rents, and (3) total firm risk.

The theoretical and managerial implications of differentiating among each of these performance facets are not trivial. Normal profits are the minimum compensation that investors require to purchase stock in a company. They equal the interest rate investors could earn in a treasury bond plus an additional risk premium (Marshall 1920). When a project generates only normal profits, its net present value (NPV) is always zero, implying that investors earn no more than what is a fair compensation for risk (Weston 1950). Undertaking only zero NPV projects might ensure a firm's long-term survival but would never confer on it the kind of competitive advantage that arises from investing in projects with "dramatic potential" (Schwartz 2006), which allows firms to thrive and grow substantially over time.

Economic rents are profits earned above those required as compensation for risk and time value of money (Weston

---

Alina B. Sorescu is Assistant Professor of Marketing and Mays Research Fellow, Mays Business School, Texas A&M University (e-mail: asorescu@tamu.edu). Jelena Spanjol is Assistant Professor of Marketing, College of Business Administration, University of Illinois at Chicago (e-mail: spanjol@uic.edu). The authors thank Ed Blair; Rajesh Chandy; Michael Hitt; Sanjay Jain; Jaideep Prabhu; William Qualls; José Antonio Rosa; Venky Shankar; Sorin Sorescu; Rajan Varadarajan; Manjit Yadav; participants at the 2006 Mays Business School Research Innovation Camp, the 2005 American Marketing Association Winter Educators' Conference, the 2006 INFORMS Marketing Science Conference; and seminar participants at University of Illinois at Chicago for many helpful comments and suggestions. They also thank Burcu Mercankaya for excellent research assistance related to data collection. They gratefully acknowledge financial support from the Mays Business School Summer Research Grant program.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

1950). Firms posting such profits are the darlings of Wall Street and the envy of their competitors. They are featured in the press, and their stock is in high demand because capital flows to these companies where it receives the highest risk-adjusted return. Thus, determinants of economic rents are much more valuable, albeit elusive, than those of normal profits; uncovering these determinants is crucial to the understanding of the true financial value of a corporate strategic action. For example, if innovation is a determinant of neither economic rents nor normal profits, managers should forsake it in favor of better investment alternatives. If innovation is a determinant of normal profits but not economic rents, managers should pursue it in the absence of any other positive NPV projects. However, if an innovation can generate economic rents, managers acting exclusively in the shareholders' interest should prioritize bringing such innovation to the market.

However, managers might also pursue other objectives. The corporate governance literature is replete with examples of divergence between the interests of shareholders and those of managers and other stakeholders (e.g., Jensen and Meckling 1976; Titman 1984). This divergence is almost always exacerbated by firm risk (Grinblatt and Titman 1998). Consider an innovation project that increases risk but generates economic rents even after accounting for this risk increase. Shareholders would welcome the project because they are more than compensated for the additional risk by an above-normal rate of return. However, the increase in risk may endanger the firm's survival prospects and have negative consequences for management, employees, and other stakeholders (Grinblatt and Titman 1998; Hirshleifer and Suh 1992). When shareholders' and managers' interests diverge, managerial decisions will account for both shareholder wealth and risk as separate dimensions of firm performance.[1]

In summary, to obtain a complete picture of innovation's impact on the entire organization, we must assess three separate aspects of firm performance: normal profits, economic rents, and risk. Consequently, we attempt to answer the following research questions:

- What is the effect of breakthrough and incremental innovations on a firm's normal profits?
- Do incremental and breakthrough innovations generate economic rents?
- What is the effect of breakthrough and incremental innovations on firm risk?

In answering these questions, we account for firm resources and marketing activities, as well as the possibility that innovation and firm value are simultaneously determined by some unobserved latent variable.[2] We empirically investigate our research questions on a data set of more than 20,000 new products, which enables us to draw conclusions about a context—consumer packaged goods (CPG)—in which the financial rewards of innovation have been understudied. Our data set contains two desirable features that enhance the generalizability and reliability of our findings: (1) To our knowledge, it is the most comprehensive database of new product introductions from firms in CPG industries, and (2) it enables us to use market-based rather than self-reported measures for our variables, thus minimizing the potential for memory and self-report biases (Golden 1992).

Our findings highlight the distinct roles of breakthrough and incremental innovations in building firm value. We find that breakthrough innovations are associated with increases in both normal profits and economic rents, whereas incremental innovations are associated only with increases in normal profits. Furthermore, breakthrough innovations are associated with increases in the risk of the innovating firm, but the higher risk is offset by above-normal stock returns. Finally, incremental innovations have no effect on the riskiness of the firm.

We organize the rest of the article as follows: The next section presents the conceptualization of breakthrough and incremental innovation in CPG industries, followed by a brief review of prior research on innovation and firm value. We then provide a description of empirical methods, discuss the results, and present our conclusions.

Conceptual Framework

Breakthrough Innovation in CPGs

In line with previous research, we define "breakthrough innovations" as new products that are the first to bring novel and significant consumer benefits to the market (e.g., Chandy and Tellis 1998; Sorescu, Chandy, and Prabhu 2003; Zhou, Yim, and Tse 2005). These benefits can range from improvements to product features, such as packaging (e.g., a new type of spout on a beverage product), to opening up an entirely new market (e.g., laundry sheets that absorb excess dye during the wash cycle). We define "incremental innovations" as new products that do not deliver novel and significant consumer benefits to the market.

Our definition of breakthrough innovation draws on prior studies that have gone beyond a product's technical capabilities in judging its innovativeness. For example, Chandy and Tellis (1998, p. 476, Table 1) and Sorescu, Chandy, and Prabhu (2003, p. 84) use the term "market breakthrough" for products that incorporate new, substantial consumer benefits. Although the literature also provides several definitions and descriptive labels for technological breakthrough innovation, we found that these definitions insufficiently capture the scope of innovation in industries in which marketing innovations may well be more prevalent than their technological counterparts.[3] To confirm the appropriateness of our definition in the CPG context, we interviewed managers from Sara Lee, PepsiCo, and Unilever and asked them to describe the features of a truly innovative product. We also examined the products that won innovativeness awards in CPG industries, such as the Spirit of Innovation Award, which is presented annually by *Prepared Foods* magazine. These industry sources consistently indicated that innovative products are those that are deemed to increase consumer benefits substantially compared with existing products, whether they do so through a new technology or a new way to package the merchandise.

---

[1] Our primary focus in this article is on total firm risk. Total risk has a systematic component that is of interest only to shareholders (Lintner 1965; Sharpe 1964; Treynor 1962), as well as an idiosyncratic component that affects not only shareholders but also managers and other firm stakeholders (Grinblatt and Titman 1998). We focus on total risk because it affects the utility of shareholders and other stakeholders alike. Thus, we use the terms "risk" and "total risk" interchangeably. As robustness checks, we also provide results from analyses on the separate systematic and idiosyncratic risk components.

[2] For example, managers may decide to innovate more when they foresee a good year for the firm.

[3] For example, in technology-intensive industries, really new products have been called "disruptive" when they abruptly change

Defining breakthrough innovation in terms of consumer benefits allows for a multidimensional view of innovativeness. New products may offer one or multiple new benefits, and these benefits may come in many forms. For example, the accessibility of a product can be increased through an innovation in its packaging (e.g., vacuum packaging eliminated the need for refrigeration), formulation (e.g., microwavable formulations simplified the preparation of certain food products), or both. In our empirical analysis, we also examine the extent to which the scope of the benefits provided by breakthrough innovations is associated with increases in firm value.

The Effect of Innovation on Firm Value

Innovation, normal profits, and economic rents. The effects of innovation on firm value uncovered by previous research range from nonexistent (Eddy and Saunders 1980) to small and accruing only under limited circumstances (Chaney, Devinney, and Winer 1991; Geroski, Machin, and Van Reenen 1993) to positive and significant (Pauwels et al. 2004; Sorescu, Chandy, and Prabhu 2003; Srinivasan et al. 2006). These apparent inconsistencies could arise from (1) differences in sample selection and definition, (2) product- rather than firm-level analysis, and (3) differences in performance metrics used.

First, many studies use samples of radical or otherwise important innovations, but they draw conclusions about innovation in general. For example, Lee and colleagues (2000, p. 25) find cumulative short-term abnormal returns of approximately 2% for a sample of innovations drawn from the *Predicast F&S Index*, selected so that "each new product introduction represented a product category that did not exist prior to the announcement date." Blundell, Griffith, and Van Reenen (1999, p. 530) find positive stock market returns for a sample of "major new technological breakthroughs." Large-scale studies on the value of incremental innovations are virtually absent, with the exception of the pharmaceutical industry (e.g., Sharma and Nelson 2004) and, more recently, the automobile industry (Srinivasan et al. 2006).

Second, most studies that examine the effect of innovation on firm value do so by assessing the stock market reaction to the introduction of selected individual products rather than to changes in the entire firm product portfolio (e.g., Chaney, Devinney, and Winer 1991; Eddy and Saunders 1980). This type of product-level analysis does not account for the trade-offs that resource-constrained firms make during the product development process, including whether to introduce fewer riskier but potentially more profitable breakthrough products or to pursue a greater number of (presumably safer) line extensions. A firm could arguably introduce two valuable breakthrough innovations but still earn overall smaller profits than it might earn from a multitude of incremental innovations that are less costly to develop and market.[4]

Finally, extant studies fail to acknowledge that the effect of innovation on firm value may markedly differ depending on the firm performance metric employed. Ignoring these differences could lead to seemingly contradictory conclusions. For example, we expect incremental innovation to be a determinant of normal profits but not economic rents. We view incremental innovation as a necessary part of any firm's normal operational strategy. However, because incremental innovations are easier to imitate, they are likely to be priced competitively and to generate no economic rents (Dixit 1980; Spence 1977).

In contrast, breakthrough innovation is neither a requisite for firm survival nor commonly part of a firm's operational strategy. Breakthrough innovations are rare events; even in high-tech industries, in which they are the bloodline of many businesses, breakthroughs represent only approximately 6% of total innovation output (Sorescu, Chandy, and Prabhu 2003). Moreover, multiple studies have identified a positive stock market reaction to the announcement of breakthrough innovations, suggesting that they often come as a pleasant surprise to investors (e.g., Lee et al. 2000; Sorescu, Chandy, and Prabhu 2003). What, then, can be said about the impact of breakthroughs on economic rents?

On the one hand, basic industrial organization theory suggests that breakthrough products will earn rents if they can be successfully differentiated from competitors' products (Tirole 1988). On the other hand, the competitive advantage conferred by product innovativeness has been called into question in two meta-analyses: VanderWerf and Mahon (1997) conclude that firms that are the first to introduce new products to market are not necessarily the most successful, and Henard and Szymanski (2001) find that product innovativeness is not a significant driver of new product performance.

The positive short-term market reaction documented around the announcement of breakthrough innovations is not necessarily an indication of economic rents. Conceivably, if breakthroughs use resources that may have served

---

the established trajectory of performance improvement in current products (Christensen and Bower 1996), "technologically discontinuous" when they dramatically advance an industry's price versus performance frontier (Anderson and Tushman 1990), or "radical technological" when they involve methods and materials that are novel to incumbents (e.g., Hill and Rothaermel 2003). Garcia and Calantone (2002) provide a review of the definitions and descriptive labels used for "really new products" in extant research.

[4] For example, Kim and Mauborgne (1997) note that 86% of new product launches are line extensions, but they account for only 62% of revenues and 39% of profits. In contrast, 14% of launches are true innovations; these generate 38% of revenues and 61% of profits. The numbers indicate that executives must balance the costs, risks, and returns of individual new products to maximize the returns to a firm's entire new product portfolio.

other productive purposes, these positive announcement returns (traditionally measured during periods lasting no more than five trading days) may be offset by a more sluggish stock market performance each time the firm is forced to forgo an otherwise lucrative investment opportunity for lack of resources. These offsetting effects, if any, will become apparent only over a longer-term measurement window. Thus, the overall long-term impact of breakthrough innovation on firm value is an empirical question.

Innovation and firm risk. Several authors have argued that breakthrough innovation is risky and have associated this risk with a lower likelihood or speed of product adoption or a lower likelihood of survival of the innovating firm. Researchers have even suggested that a "curse of innovation" can afflict firms that pursue breakthrough products; these products often fail in the marketplace because managers consistently overvalue their benefits relative to existing products, whereas consumers systematically undervalue them or forgo them altogether in favor of more familiar products (Gatignon and Robertson 1985; Gourville 2005). Furthermore, in markets started by a radical product, the first firms to enter the market have the lowest survival probability (Min, Kalwani, and Robinson 2006). Perhaps consistent with this view, the numbers of breakthroughs launched in recent years have indeed been declining.[5]

However, although it has been argued that breakthrough innovations are risky, to our knowledge, there is no empirical research that measures the effect of innovation on firm risk. Indeed, research showing the impact of any type of marketing strategy on financial risk is scarce (cf. Madden, Fehle, and Fournier 2006; McAlister, Srinivasan, and Kim 2007; Singh, Faircloth, and Nejadmalayeri 2005). A possible reason is that the measures of shareholder value most commonly used in research—Tobin's q and abnormal stock returns—already incorporate the effects of risk on shareholder wealth. Specifically, abnormal returns are those beyond what is required to compensate investors for risk, and the numerator in the Tobin's q ratio is the firm's market value, which is itself the sum of future cash flows discounted at the risk-adjusted cost of equity.

However, neither Tobin's q nor abnormal returns can accurately capture the effect of risk on firms' survival prospects or on the utility of stakeholders other than shareholders. Specifically, total firm risk is inversely related to the firm's survival probability, so firms facing high risk levels have greater probabilities of becoming financially distressed and filing for bankruptcy protection. The consequences of financial distress are almost always unfavorable for creditors, management, customers, and employees. Thus, the utility of these stakeholders is decreasing in total firm risk (Hirshleifer and Suh 1992; Jensen and Meckling 1976). Moreover, unlike shareholders who are automatically compensated for risk in the form of higher rates of return, in general, these other stakeholders (with the possible exception of creditors) receive no such compensation and are typically worse off when total risk increases, all else being equal (Grinblatt and Titman 1998; Maximovic and Titman 1991). Consequently, examining risk is of sufficient independent interest and is a necessary step to obtaining a complete picture of how the entire organization responds to the firm's innovative strategy.

In summary, there is mixed evidence on the potential effect of innovation on firm risk. On the one hand, there is evidence that breakthrough innovations have a high rate of failure (e.g., Neff 2005). Firms that pursue them may increase the volatility of their future cash flows, leading to a higher volatility in stock prices and, thus, higher risk (Grinblatt and Titman 1998). Moreover, pursuing breakthroughs may take away resources from other, less risky projects, further increasing cash flow volatility. On the other hand, breakthrough innovations are unique products that may, at least for a while, enjoy a de facto monopoly protection, either because of patents or because they are difficult to imitate (Tirole 1988). After they are launched and adopted, they could conceivably increase the predictability of future cash flows because of the lack of close substitutes. Thus, the effect of breakthrough and incremental innovation on firm risk is an empirical question.

Research Design

Empirical Context

We test our research questions using data from CPG industries. We chose this context for several reasons. Consumer packaged goods account for a substantial portion of the U.S. economy. In 2005, U.S. spending on CPGs exceeded $2 trillion (Bureau of Economic Analysis 2007). Nevertheless, despite their economic and social importance, these industries are currently underrepresented in empirical research on financial rewards to innovation. Knowledge in this area is based on generalizations from high-tech industries (e.g., Bayus, Erickson, and Jacobson 2003; Blundell, Griffith, and Van Reenen 1999), despite the idiosyncratic characteristics of high-tech markets, such as a higher rate of technological innovation or greater network externalities.

Data and Sample

We construct our data set after processing and aligning data from multiple sources. Our long-term abnormal return metrics (we discuss these subsequently) require the use of benchmark portfolios that provide an adjustment for size, book-to-market, and momentum (Daniel et al. 1997; Wermers 2003). We obtain data on these benchmarks from Russ Wermers's Web site at the University of Maryland (http://www.smith.umd.edu/faculty/rwermers/ftpsite/Dgtw/coverpage.htm). We obtain firm-level accounting data, as well as data on the determinants of innovation and performance, from COMPUSTAT, Schonfeld reports, and SDC Platinum. Innovation data come from Productscan, which is a subscription-based database that has tracked CPG introductions since the early 1980s. Productscan compiles an extensive database through constant interactions with manufacturers. It also scans trade publications, visits trade conferences, and regularly examines store shelves across the country. The complete list of the variables used in the analysis, their operationalizations, and data sources appear in Table 1.

---

[5] According to a recent study conducted by the Product Development and Management Association, both new-to-the world and new-to-the-company innovations decreased dramatically between 1990 and 2004 (44% and 30%, respectively); nonbreakthrough innovation made up 89% of new product development portfolios across firms in 2004 (Cooper 2005).

Each product listed in Productscan has a detailed record that includes a wealth of information, including the name of the product, its manufacturer, and the manufacturer's corporate affiliate; a description of product claims; the product category; flavor (when relevant); ingredients; package and shelving type; and, in many cases, even a picture of the product. These records also include a date of introduction, which is either the product's launch announcement date or the date on which the Productscan editors first became aware of the product's launch. Finally, when a product is added to the database, it is assigned an innovativeness rating based on whether it is innovative in one or more of the following areas: positioning, packaging, new markets, merchandising, formulation, and technology. We use Productscan's ratings to classify new product introductions as breakthrough or incremental. The complete definitions of Productscan's innovativeness dimensions appear in Table 2, along with illustrative examples drawn from our sample.

To construct our sample, we first select from the Center for Research on Security Prices (CRSP) all publicly traded

TABLE 1
Variables and Data Sources

Conceptual Variable	Notation	Measured Variable	Data Source
Dependent Variables: Performance and Risk			
Firm value	Tobin's $q_{i,t}$	Tobin's q	CRSP, COMPUSTAT
Abnormal returns	$BHAR_{i,t}$	Benchmark-adjusted buy-and-hold abnormal returns	CRSP, Russ Wermers's Web site
Firm risk	$Risk_{i,t}$	Standard deviation of stock returns	CRSP
Independent Variables: Innovation			
Incremental innovation	Incremental $innovation_{i,t}$	Annual firm-level count of products classified by Productscan as incremental	Productscan
Breakthrough innovation	Breakthrough $innovation_{i,t}$	Annual firm-level count of products classified by Productscan as offering a substantially new consumer benefit	Productscan
Unanticipated incremental innovation	Unanticipated incremental $innovation_{i,t}$	Residuals from an autoregressive model of the incremental innovation counts	Productscan
Unanticipated breakthrough innovation	Unanticipated breakthrough $innovation_{i,t}$	Residuals from an autoregressive model of the breakthrough innovation counts	Productscan
Control Variables			
Firm size	$Assets_{i,t}$	Natural logarithm of total assets	COMPUSTAT
Organizational slack	Organizational $slack_{i,t}$	Ratio of net cash flows from operating activities to total assets	COMPUSTAT
Fixed asset intensity	Fixed asset $intensity_{i,t}$	Ratio of fixed assets to total firm assets	COMPUSTAT
Advertising	$Advertising_{i,t}$	Advertising expenditures	COMPUSTAT and Schonfeld Reports
Acquisitions	$Acquisitions_{i,t}$	Annual count of acquisitions undertaken by parent firm	SDC Platinum and LexisNexis
Performance	$ROA_{i,t}$	Return on assets	COMPUSTAT

TABLE 2
Productscan Innovativeness Dimensions and Illustrative Examples

Innovative Dimension	Product Example	Year of Introduction	Claim
Positioning the product to new users or usage (positioning)	Ultra Gain laundry detergent—with odor removing bleach	1993	First detergent marketed as capable of removing the odor from dirty laundry.
Providing a consumer benefit with new packaging (packaging benefit)	Yoplait Go-Gurt	1998	Yogurt packaged in a tube that eliminates the need for a spoon or refrigeration.
Opening up a new market for the product (new market)	McCormick Garden Fare Freezer Jam gelling powder	1993	Creates a new market for homemade sugar-free jam by allowing customers to make freezer jam with a higher fruit content and less sugar than the pectin method.
Merchandising	Kellogg's Breakfast Mates	1998	Ready-to-eat, single-serve combination packs that pair cereal with milk (and spoon), available in the refrigerated case.
Offering additional value through a new formulation (formulation)	Tofutti Better Than Cream Cheese	1988	Soy-based product similar in taste and texture to traditional cream cheese that is milk and butterfat free and contains no cholesterol.
Introducing new technology to the market (technology)	Glad Press 'n Seal plastic food wrap	2003	First wrap that seals to various surfaces, including paper, plastic, and wood. It uses Griptex, a proprietary gripping technology.

firms whose primary Standard Industrial Classification codes are among those categorized by Productscan. We identify 308 U.S.-based firms that have been publicly traded for at least 12 months during the period we study (i.e., 1985–2003). We search the Productscan database using the name of each relevant CRSP firm and download all innovation records corresponding to these firms. We perform multiple data checks to ensure that the innovation activity of firms in our sample is accurately represented. For example, Productscan includes some duplicate entries of innovations, in which a product is first recorded when the Productscan editors become aware of it and then is recorded again when more information about that product becomes available. We eliminate such updates from the database. We also collect data on all mergers and acquisitions in which firms in our sample were targets. We mark all product records downloaded from Productscan that could have been affected by such acquisitions, and we check that they are indeed assigned to the correct firm (i.e., to the firm that actually introduced it). We include in our data set all new products introduced by each firm, regardless of whether they were invented in-house or developed from the pipeline of an acquired firm, but we control for acquisitions as a potential determinant of innovation output. Our final sample contains 153 publicly traded firms that introduced at least one new product between 1985 and 2003, for a total sample of 22,532 new product introductions.

One advantage of our data set is that it does not suffer from survival bias. The data set includes a large sample of CPG publicly traded firms, regardless of their performance or eventual survival. It also contains all new products introduced by these firms, regardless of their eventual success in the market. A second advantage of our data set is that product introductions are recorded contemporaneously (rather than subsequently to introduction), ensuring that a potential memory bias does not affect which products are included in the sample.

Model Formulation

Our objective is to examine the effect of incremental and breakthrough innovation on firm performance and risk. We begin with a discussion on how to relate the economic concepts of normal profits and economic rents to two empirical measurements of firm value: Tobin's q and abnormal returns. We then present our main equations that model the determinants of Tobin's q, abnormal returns, and risk. We follow with a discussion of the estimation issues associated with these equations and propose a two-stage instrumental variables method to address the possibility that the innovation output variables are correlated with the error term. We continue with a presentation of the first-stage equations required by this instrumental variables procedure, which are models of the determinants of breakthrough and incremental innovation. We end with a detailed explanation of how we operationalize each variable used in the analysis. In each case, subscript "i" represents the firm and subscript "t" represents the calendar year.

Relationship among normal profits, economic rents, Tobin's q, and abnormal returns. We do not directly observe economic rents or normal profits. However, we observe two related measures of firm value: Tobin's q and abnormal returns. Using a simple, two-period analytical model, we can show that if innovation earns economic rents, its relationship with Tobin's q is positive.[6] However, the relationship between innovation and Tobin's q is positive even if innovation earns only normal profits. In contrast, inferences obtained from abnormal returns are always unambiguous. The relationship between abnormal returns and innovation is positive if and only if innovation earns economic rents, and it is zero if and only if innovation earns only normal profits. Likewise, this relationship is negative if and only if innovation earns below-normal profits that are destructive of firm value.[7]

---

[6] Additional details are available on request.

Determinants of Tobin's q, abnormal returns, and risk. We estimate the following two models:

$$
\begin{aligned}
(1)\quad \text{Tobin's } q_{i,t} = {} & \alpha_{10} + \beta_{11}\text{Incremental innovation}_{i,t} \\
& + \beta_{12}\text{Breakthrough innovation}_{i,t} \\
& + \beta_{13}(\text{Incremental innovation}_{i,t} \\
& \times \text{Breakthrough innovation}_{i,t}) \\
& + \beta_{14}\text{Advertising}_{i,t} + \beta_{15}\text{Firm size}_{i,t} + \varepsilon_{1i,t}, \text{ and}
\end{aligned}
$$

$$
\begin{aligned}
(2)\quad \text{Abnormal returns}_{i,t} = {} & \alpha_{20} + \beta_{21}\text{Unanticipated incremental} \\
& \quad \text{innovation}_{i,t} \\
& + \beta_{22}\text{Unanticipated breakthrough} \\
& \quad \text{innovation}_{i,t} \\
& + \beta_{23}\text{Unanticipated (Incremental} \\
& \quad \text{innovation}_{i,t} \times \text{Breakthrough} \\
& \quad \text{innovation}_{i,t}) \\
& + \beta_{24}\text{Unanticipated changes in} \\
& \quad \text{advertising}_{i,t} \\
& + \beta_{25}\text{Unanticipated changes in ROA}_{i,t} \\
& + \varepsilon_{2i,t}.
\end{aligned}
$$

Our third model measures the effect of innovation on firm risk:

$$
\begin{aligned}
(3)\quad \text{Risk}_{i,t} = {} & \alpha_{30} + \beta_{31}\text{Incremental innovation}_{i,t} \\
& + \beta_{32}\text{Breakthrough innovation}_{i,t} \\
& + \beta_{33}(\text{Incremental innovation}_{i,t} \\
& \times \text{Breakthrough innovation}_{i,t}) \\
& + \beta_{34}\text{Advertising}_{i,t} + \beta_{35}\text{ROA}_{i,t} + \varepsilon_{3i,t}.
\end{aligned}
$$

Because we do not know the exact functional form of the relationship among innovation, firm value, and risk, we also estimate nonlinear versions of Equations 1–3 that contain squared terms for the innovation variables.

The choice of the determinants of firm value and risk in Equations 1–3 stems from previous literature. For example, advertising expenditures and firm size have been traditionally included in models that assess the relationship between innovation and firm value (e.g., Pauwels et al. 2004; Srinivasan et al. 2006). Changes in return on assets (ROA) have been shown to be associated with changes in abnormal returns (Aaker and Jacobson 1994). Finally, we acknowledge that other variables, such as pricing strategy and the magnitude of a firm's distribution system, could also influence the relationship between innovation and firm value, but data limitations prevent us from including such variables in our equations. However, excluding these variables does not change our conclusion as long as the excluded variables are uncorrelated with innovation output.

We include year dummies in all three models, and when necessary, we adjust standard errors to account for heteroskedasticity according to White's (1980) method. The Wooldridge test for autocorrelation in panel data indicates the presence of first-order autocorrelation in Equations 1 and 3 when these equations are estimated with random-effects generalized least squares models. To correct for autocorrelation, we estimate Equations 1 and 3 using a feasible generalized least squares procedure that allows for the coefficient of the AR(1) process to be specific to each panel (Beck and Katz 1995).

Equation 2 requires a different estimation approach. Abnormal returns capture only new, unanticipated information, the effect of which is not already incorporated into stock prices (Aaker and Jacobson 1994; Sorescu, Chandy, and Prabhu 2007). In other words, abnormal returns are positive only if (1) innovation is positively related to firm value and (2) the firm innovates more than expected. This requires the use of unanticipated components of innovation on the right-hand side of Equation 2. As in the work of Aaker and Jacobson (1994), we estimate the unanticipated components of breakthrough innovation, incremental innovation, advertising, and ROA as residuals from first-order autoregressive models. The coefficients of the independent variables in Equation 2 show the extent to which unanticipated changes in innovation, advertising, and firm performance lead to unexpected changes in stock prices. Finally, all fixed effects are already reflected in previous period stock prices and thus are accounted for in the abnormal stock return measures (Sorescu, Chandy, and Prabhu 2007).

Time frame for measuring performance. We conduct annual, firm-level analyses. We measure risk and abnormal returns, which are flow variables, over annual windows, from the first trading day in January to the last trading day in December each year. Tobin's q is a stock variable, and we measure it at the end of each calendar year. As the main time unit for our analysis, we chose the calendar year instead of specific event dates. We do this for two reasons.

First, many new CPGs are introduced without being announced, which is why Productscan representatives regularly examine retailers' shelves for new products to be included in their database. Firms in our data set introduced as many as 346 innovations per year. Such firms are unlikely to set up press conferences to announce each introduction, and investors are not able to take note of every market introduction. However, investors will eventually see summary activity reports or other cumulative evidence of a stream of innovations and will update stock prices from the insights gleaned about the strategic orientation of the firm. Measuring stock market reaction over an entire year in response to firms' aggregate innovation output enables us to incorporate the price effect of the information flow that occurs throughout the year.

---

[7] These three cases correspond to situations in which the NPV of the innovation project is, respectively, positive, zero, and negative.

Second, several factors that are crucial to the performance of a new product are not readily quantifiable at the time of introduction. For example, the true functional performance of the product may not match the initial product claims, subsequently diminishing the expected value of cash flows from innovation. These cash flows may also be significantly affected by unanticipated competitive reactions. Therefore, investors may need some time to learn about the financial value of innovations before they gradually incorporate their updated expectations into stock prices (e.g., Pauwels et al. 2004; Sorescu, Shankar, and Kushwaha 2007).

Accounting for endogeneity. A possible concern in estimating Equations 1 and 3 is that the error terms might be correlated with the independent variables of interest (breakthrough and incremental innovation).[8] In the case of Equation 1, this could occur if an unexpected event simultaneously affects both firm value and the likelihood to introduce new products. For example, management's propensity to innovate might be greater when there is an unexpectedly high end-of-year firm value (Tobin's q), or the propensity to innovate might be driven by unexpected changes in macroeconomic variables that also affect firm value (e.g., an unexpected decline in interest rates).

In the case of Equation 3, the correlation between innovation variables and $\varepsilon_{3i,t}$ could occur if management decided to introduce more new products after observing unexpectedly high volatility during current year (t), perhaps attempting to stabilize the company's cash flows. Specifically, we can view $\varepsilon_{1i,t}$ and $\varepsilon_{3i,t}$ as the sum of two terms:

$$\varepsilon_{1i,t} = u_{1i,t} + v_{1i,t}, \text{ and } \varepsilon_{3i,t} = u_{3i,t} + v_{3i,t},$$

where $u_{1i,t}$ and $u_{3i,t}$ are uncorrelated with Incremental innovation$_{i,t}$ and Breakthrough innovation$_{i,t}$ and $v_{1i,t}$ and $v_{3i,t}$ are correlated with the two innovation variables. The $v_{1i,t}$ ($v_{3i,t}$) term captures the effect of all latent variables that could simultaneously affect Tobin's q (risk) and the innovation variables.

A way to correct for this potential problem would be to include the latent variables themselves in the right-hand side of Equations 1 and 3. However, because we cannot possibly identify all potential latent variables, we follow the more common approach and choose instead to instrument Incremental innovation$_{i,t}$ and Breakthrough innovation$_{i,t}$. When we choose our instruments, we must ensure that they are correlated with the actual variables but uncorrelated with latent variables $v_{1i,t}$ and $v_{3i,t}$, respectively.

Previous research has suggested that current innovation is related to Innovation$_{i,t-1}$, Firm size$_{i,t-1}$, Organizational slack$_{i,t-1}$, Fixed asset intensity$_{i,t-1}$, and Acquisitions$_{i,t-1}$ (see, e.g., Hambrick and McMillan 1985; Herold, Jayaraman, and Narayanaswamy 2003; Sorescu, Chandy, and Prabhu 2003, 2007). We refer to these five variables as "determinants" of innovation. A notable property of these determinants is that they are not only correlated with the innovation variables but also uncorrelated with $v_{i,t}$. The reason is that the determinants are lagged by one year (t – 1), whereas the latent variables $v_{i,t}$ capture unexpected occurrences during the current year (t). Given this property, we can obtain instruments for breakthrough innovation and incremental innovation by using the information contained in all five determinants (see, e.g., Maddala 1992, p. 373). These instruments are the predicted values from the following models:[9]

$$\begin{aligned}(4)\quad \text{Incremental innovation}_{i,t} = {} & \alpha_{40} \\ & + \beta_{41}\text{Incremental innovation}_{i,t-1} \\ & + \beta_{42}\text{Firm size}_{i,t-1} \\ & + \beta_{43}\text{Organizational slack}_{i,t-1} \\ & + \beta_{44}\text{Fixed asset intensity}_{i,t-1} \\ & + \beta_{45}\text{Acquisitions}_{i,t-1} + \varepsilon_{4i,t}, \text{ and}\end{aligned}$$

$$\begin{aligned}(5)\quad \text{Breakthrough innovation}_{i,t} = {} & \alpha_{50} + \beta_{51}\text{Breakthrough} \\ & \quad \text{innovation}_{i,t-1} \\ & + \beta_{52}\text{Firm size}_{i,t-1} \\ & + \beta_{53}\text{Organizational slack}_{i,t-1} \\ & + \beta_{54}\text{Fixed asset intensity}_{i,t-1} \\ & + \beta_{55}\text{Acquisitions}_{i,t-1} + \varepsilon_{5i,t}.\end{aligned}$$

Because the dependent variables in Equations 4 and 5 are count variables, we cannot estimate these models with ordinary least squares. Instead, we estimate them using a zero-inflated count negative binomial model (Long and Freese 2003). Including previous innovation and several firm-level control variables also ensures that firm-specific effects are adequately accounted for in the model. Consistent with the instrumental variables procedure, all right-hand-side variables in Equations 4 and 5 are exogenous to our system of equations.[10] The Appendix provides additional details of this instrumental variables procedure.

---

[8] This concern does not apply to Equation 2, in which the independent variables represent unexpected innovation activity. Our estimation approach of Equation 2 follows that of Aaker and Jacobson (1994).

[9] Equations 1 and 3 also include a nonlinear term equal to the product of incremental and breakthrough innovations. To address this nonlinearity, we follow the procedure that Kelejian (1971) proposes. Rather than multiplying the predicted values of incremental and breakthrough innovations, we first multiply the actual values and then estimate a predicted term for the product. The predicted term serves as an instrument for the interactive term in Equations 1 and 3.

[10] We also considered research and development (R&D) intensity a potential determinant of innovation, but because of data limitations, we did not include it in our main model. Data on R&D expenditure are available for less than 60% of our sample. To preserve sample size and because the coefficient of R&D is not significant in the subsample for which data are available, we exclude R&D from our main model. The size of category margins could also potentially affect innovation. Unfortunately, data on category margins are not publicly available. Not including this variable in the model is a limitation of this study that we hope will be addressed in further research.

Dependent Variables

Tobin's q. Following the work of Kaplan and Zingales (1997) and Gompers, Ishii, and Metrick (2003), we compute year-end measures of Tobin's q as the ratio of the market value to book value of firm assets. We estimate the market value of assets as the book value of assets plus the market value of common stock (obtained from CRSP), less the book value of common stock, less the amount of deferred taxes.

Buy-and-hold abnormal returns. To compute stock abnormal returns, we use the measure of buy-and-hold abnormal returns (BHAR), as defined by Barber and Lyon (1997). This compares the actual return the innovative firm earns with a benchmark return the firm would have earned without any unexpected product introductions. Following Daniel and colleagues (1997), we use as benchmark a portfolio of stocks that belong to the same size, book-to-market, and momentum quintiles as the stock of the firm under consideration. This is because prior research shows that these three characteristics most accurately describe the cross-section of stock returns (Fama and French 1992; Jegadeesh and Titman 1993). Specifically, we compute BHARs as follows:

$$\text{BHAR}_{i,t} = \prod_{m=1}^{m=12}(1+R_{i,m,t}) - \prod_{m=1}^{m=12}(1+R_{j(i,m,t)}),$$

where t is the calendar year of interest; m is the calendar month within each year; $R_{i,m,t}$ is the rate of return of innovating firm i in month m of year t; and $R_{j(i,,m,t)}$ is the return of a control portfolio j that includes all stocks that belong to the same size, book-to-market, and momentum quintiles as firm i at the beginning of month m of year t.[11]

Risk. Under the standard assumptions of Von Neumann and Morgenstern's (1944) expected utility theory, when asset returns are normally distributed, the total risk of an asset is the volatility (or standard deviation) of its return distribution, which in turn reflects the volatility of its future cash flows (Copeland and Weston 1992). To obtain annual measures of firm risk, we compute the standard deviation of daily stock returns each calendar year (Ronn and Verma 1986).

Independent Variables

Incremental innovation. We construct annual measures of firm-level incremental innovation by summing the number of incremental new products that each firm introduced each year during the 1985–2003 period. As we noted previously, Productscan considers a product incremental if it does not offer any significantly new consumer benefits. For example, although Tropicana's fruit juice ice-cream bar constituted an entry into a new product category for Tropicana, it was not a breakthrough innovation in terms of the benefits offered to consumers, given that other brands of frozen fruit juice bars already existed on the market.

Breakthrough innovation. We construct annual measures of firm-level breakthrough innovation by summing the number of breakthrough new products each firm introduced each year during the 1985–2003 period. We classify breakthroughs as products that Productscan recognizes at the time of introduction as industry firsts on one or more of the six innovativeness dimensions in Table 2. We subsequently report robustness tests conducted using the individual innovativeness dimensions separately.

In adopting Productscan's innovativeness classification, we realize two significant advantages over existing alternatives. First, we evaluate innovativeness from the perspective of the industry, not the company. Stock market analysts and investors tracking CPG companies adopt an industry perspective because industrywide breakthroughs, rather than new-to-the-firm products, are the most likely to generate economic rents and to increase stock returns (Tirole 1988). Second, the classification we use is made at the time the product enters the market, which eliminates possible temporal memory biases (Golden 1992). An example of a breakthrough innovation is Clorox Anywhere Hard Surface spray, a sanitizing spray that is as gentle as water but kills common household bacteria found on hard surfaces.

It could be argued that in high-tech industries, innovations in packaging or merchandising would not be considered sufficiently pathbreaking to warrant labeling the product a breakthrough. However, this does not appear to be the case in CPGs. For example, at the 2005 Spirit of Innovation Awards, a food industry panel awarded DiGiorno Microwave Pizza first place, Pringles Prints (potato chips with trivia printed on them) second place, and Oh Boy! Oberto Beef Jerky Crisps third place. An industry outsider may not consider these products breakthroughs, but such innovations can be the key to healthy growth and profitability for a CPG company. Indeed, Go-Gurt, the yogurt in a tube dubbed the "innovation that gave yogurt an entertaining way into kids' lunch boxes," generated $124 million in sales for General Mills (*Advertising Age* 2005). Furthermore, that a merchandising innovation, such as Pringles Prints, and a formulation innovation, such as DiGiorno Microwave Pizza, are recognized as the most innovative products in the food industry adds validity to Productscan's innovativeness classification.

Control Variables

The following control variables are used in the first-stage models of innovation counts:

- Firm size$_{i,t}$: The larger the firm, the more innovations it is expected to introduce (Pauwels et al. 2004; Sorescu, Chandy, and Prabhu 2003). We measure firm size using the natural logarithm of the book value of assets (e.g., McAlister, Srinivasan, and Kim 2007).
- Organizational slack$_{i,t}$: The greater the organizational slack, the more internal resources can be dedicated to innovation, and the more likely it is that the firm will innovate. Organizational slack in the form of available, excess resources beyond those essential for ongoing operations is necessary for firms to pursue experimentation because it effectively shelters them from the uncertainties associated with innovation projects (Cyert and March 1963; Herold, Jayaraman, and Narayanaswamy 2003). We measure organizational slack using the

---

[11] For complete details about the construction of these control portfolios, see the Appendix in Daniel and colleagues (1997).

ratio of net cash flows from operating activities to total firm assets (Davis and Stout 1992).

- Fixed asset intensity$_{i,t}$: Fixed asset intensity assesses the level of cash and other liquid assets available to flexibly finance innovation projects. The higher a firm's ratio of fixed assets to total assets, the lower are the internal resources that can be dedicated to innovation, and the less likely it is that the firm will innovate (e.g., Hambrick and McMillan 1985; Miles and Snow 1978).
- Acquisitions$_{i,t}$: Acquisitions could increase the innovative output of firms because some targets are often acquired for their new product pipeline (Sorescu, Chandy, and Prabhu 2007). Conversely, acquisitions could reduce the amount of resources available for innovation, making it less likely for firms to innovate. We account for the number of acquisitions firms undertake each year as a potential determinant of innovation output in the following year.

In the second-stage Equations 1–3, we use one or more of the following variables:

- Advertising$_{i,t}$: Advertising expenditures have been shown to enhance the financial performance of the firm (Joshi and Hanssens 2004) and to reduce systematic risk (McAlister, Srinivasan, and Kim 2007). We measure advertising using advertising expenditures obtained from COMPUSTAT and Schonfeld reports.
- Performance$_{i,t}$: We obtain returns on assets (ROA$_{i,t}$) from COMPUSTAT.
- Unanticipated change in performance$_{i,t}$: We compute a measure of unanticipated change in ROA$_{i,t}$ as in the work of Aaker and Jacobson (1994).

Results

Descriptive Statistics

Table 3 presents descriptive statistics on the type of CPG innovations introduced between 1985 and 2003. Of the 22,532 products in our sample, only 1567 (approximately 7%) are breakthrough new products that bring at least one significantly new benefit to the market. They were introduced by 95 of the 153 firms in our sample. A large number of breakthroughs (966) involved a new formulation, but only 48 introduced a brand-new technology. Examples of these rare technological breakthroughs include Simply Eggs (a low-cholesterol liquid whole egg), the Swiffer WetJet cleaning system, and the Mr. Clean AutoDry Car Wash System. Although the percentage of breakthrough to total innovations (7%) appears to be in line with high-tech industries, the percentage of technological breakthroughs to total innovations (.2%) is significantly lower, further highlighting the need for empirical insights specific to the CPG industry.[12] The low incidence of technological breakthroughs further justifies the use of a breakthrough definition that is based on the newness and significance of consumer benefits rather than on the technological newness of the product.

Table 3 also presents descriptive statistics on the dependent and control variables. The sample offers substantial variation in size, performance, and firm resources. Moreover, the mean value of the BHARs is not significantly different from zero, indicating that, performance-wise, this appears to be a representative sample of publicly traded companies. Table 4 presents correlations between the variables used in the study.

Determinants of Performance and Risk

The effect of incremental and breakthrough innovations on normal profits and economic rents. The results obtained from estimating Equation 1 appear in the top section of Panel A in Table 5. The top line presents the results for the linear specification. Both incremental (β_{11} = .00075, $p < .01$) and breakthrough (β_{12} = .034, $p < .05$) innovations are significant determinants of Tobin's q, after we account for the effects of firm size and advertising. Their estimated coefficients show that, all else being equal, each additional incremental innovation is associated with a .075% increase in Tobin's q. In the case of breakthroughs, this coefficient is 3.4%, almost 50 times higher. The second line presents the results for the nonlinear specification. For both types of innovation, the coefficients on the linear terms are significant and positive ($p < .01$), whereas those on the squared terms are significant and negative ($p < .01$). This means that the relationship between innovation and firm value is increasing at a decreasing rate.[13]

As we mentioned previously, a positive relationship between Tobin's q and the innovation variables is consistent with the existence of either economic rents or normal profits. To discriminate between the two, we must examine the coefficients obtained in Equation 2, in which BHARs are used as measures of firm performance. These results appear in the middle section of Table 5, Panel A. Contrary to the Tobin's q results, we find that incremental innovation has no effect on BHARs, but the effect of breakthroughs is significant (β_{22} = .00085, $p < .05$). The results from the nonlinear specification show again that the relationship between BHARs and breakthrough innovation increases at a decreasing rate.

The significant, positive relationship between BHARs and breakthrough innovation is evidence that breakthroughs earn economic rents. The intuition is that with each unexpected breakthrough introduction, the market value of the firm goes up by an amount equal to the NPV of the project,

---

[12] In the pharmaceutical industry, approximately 6% of all innovations introduced in 1992–2002 were radical innovations (Sorescu, Chandy, and Prabhu 2003). This statistic is also in line with other research in which the percentage of nonbreakthrough innovations was found to be 89% (Cooper 2005).

[13] A possible concern with this interpretation is that the relationship could have an inverse U shape. However, further analysis reveals that this is not the case. The first derivative of q with respect to breakthrough innovation is zero when the value of breakthrough innovation is 24 units per year. Thus, if the relationship had an inverse U shape, the "negative slope" portion would occur to the right of the 24-units point. However, as Table 3 shows, the breakthrough innovation variable in our sample never takes on a value greater than 24. This implies that the relationship between innovation and firm value monotonically increases within the range of values in our sample. A similar analysis (and conclusion) applies to the incremental innovation variable and also to the significant nonlinear terms in the BHAR and risk equations. In all cases, the negative-slope portion of the relationship occurs out of sample

TABLE 3
Descriptive Statistics

Innovation

	Total Across the Sample	M[a]	SD[a]	Minimum[b]	Mdn[c]	Maximum[d]	Skewness[c]
Incremental Innovation	20,965	9.32	5.31	0	1	327	5.81
Breakthrough Innovation[f]	1567	1.13	1.07	0	0	24	4.77
Formulation	966	.74	.88	0	0	15	3.87
Merchandising	35	.15	.42	0	0	4	4.43
New market	28	.23	.37	0	0	2	2.80
Packaging benefit	407	.48	.75	0	0	3	7.72
Positioning	480	.46	.63	0	0	9	3.67
Technology	48	.16	.41	0	0	4	4.34

Performance

	Sample Size	M[a]	SD[a]	Minimum[b]	Mdn[c]	Maximum[d]	Skewness[c]
Tobin's q	1827	1.89	.65	.26	1.57	13.21	2.62
BHARs[g]	1838	–.005	.03	–.26	–.002	.26	–.32
Risk[h]	1958	.03	.01	.004	.02	.21	2.48

Control Variables

	Sample Size	M[a]	SD[a]	Minimum[b]	Mdn[c]	Maximum[d]	Skewness[c]
Assets (millions)	1961	1,960	767	.427	276	96,175	5.91
Organizational slack	1648	.10	.05	0	.10	.68	1.37
Fixed asset intensity	1945	.59	.12	0	.58	2.59	.85
Acquisitions	1737	.53	.75	0	0	13.00	3.97
Advertising (millions)	1586	216.9	60.5	0	159.6	1492	1.47
ROA (%)	1998	2.87	8.49	–202.99	5.80	132.78	–4.16

[a] Average of the annual, firm-level values (mean and standard deviation, respectively) for all years with data availability.

[b] Minimum of all annual, firm-level values.

[c] Based on all annual, firm-level values.

[d] Maximum of all annual, firm-level values.

[e] Sample size refers to the combination of firm/years for which data are available. Annual innovation counts are available for 1993 firm-year observations.

[f] Overall breakthrough innovations include all products that were innovative on one or more dimensions. For example, a breakthrough product classified as incorporating a new technology and a substantially new packaging benefit would be counted in both the technology and packaging counts of innovation presented in this table. As such, 1209 of the 1567 products were innovative on only one of the six dimensions, whereas only 358 were innovative on two or more dimensions.

[g] BHARs are expressed in annual terms, decimal point notation. Thus, the mean BHAR is –.5% per year

[h] Risk is expressed in daily terms, decimal point notation. Thus, the mean risk is 3% per day, which corresponds to 47% per year. (Assuming 250 trading days in a year, the annual standard deviation is $3\% \times \sqrt{250} = 47\%$.)

and the NPV is itself defined as the present value of all future economic rents (Sorescu, Chandy, and Prabhu 2003). Conversely, a project with zero NPV earns no rents and has no unexpected impact on firm value. This is the case with incremental innovation, whose coefficient β_{21} in Equation 2 is not only nonsignificant but also negative.

Taken together, the results from the Tobin's q and abnormal returns analyses suggest that incremental innovation is associated with increases in normal profits but not economic rents. However, breakthrough innovation is associated with increases in both normal profits and economic rents.

The effect of incremental and breakthrough innovations on total firm risk. The results on risk (Equation 3), which appear in the bottom section of Panel A in Table 5, complement the findings we obtained with the BHARs and Tobin's q measures. We find that incremental innovation has no effect on firm risk. However, breakthrough innovation is positively associated with increases in risk ($\beta_{32} = .00075$, $p < .01$). This may explain why some managers have reduced their emphasis on highly innovative products (Cooper 2005). One possibility is that managers are unaware that breakthrough innovations earn economic rents despite the increase in firm risk. Another possibility is that

TABLE 4
Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11
1. Incremental Innovation	1.00										
2. Breakthrough Innovation	.71	1.00									
3. Tobin's q	.24	.19	1.00								
4. BHARs	.08	07	.18	1.00							
5. Risk	–.21	–.22	–.07	–.31	1.00						
6. Log of assets	.46	.45	.09	.15	–.62	1.00					
7. Organizational slack	.14	.15	.44	.21	–.32	.27	1.00				
8. Fixed asset intensity	–.09	–.06	–.14	–.03	–.07	.07	.13	1.00			
9. Acquisitions	.09	.11	.01	.05	–.16	.28	.01	–.07	1.00		
10. Advertising	.19	.21	.14	.08	–.28	.36	.16	–.02	.08	1.00	
11. ROA	.13	.13	.22	.14	–.42	.30	.44	.01	.07	.17	1.00

Notes: All values greater than .06 (in absolute value) are significantly different from zero at the $p < .01$ level, whereas those greater than 05 are significantly different from zero at the $p <$ 05 level.

managers may be concerned that the increase in risk could have negative consequences on their own employment prospects, as well as the welfare of other firm stakeholders, despite the obvious benefit for shareholders (Hirshleifer and Suh 1992). We return to these arguments in the "Summary and Discussion" section.

Finally, and not surprisingly, we find that firm size ($\beta_{35} = -.0058$, $p < .01$) and ROA ($\beta_{36} = -.00026$, $p < .01$) are both negatively related to firm risk. The results of the nonlinear specification demonstrate again that the relationship between risk and breakthrough innovation increases at a decreasing rate. Panel B of Table 5 summarizes the effects of breakthrough and incremental innovation on the three measures of firm performance.

Determinants of Innovation

We now present the first-stage results of the instrumental variables method previously described. Specifically, we examine the determinants of breakthrough and incremental innovation, whose predicted values are used as instruments in Equations 1 and 3. The results appear in Panel A of Table 6.

We estimate Equations 4 and 5 using a zero-inflated negative binomial model because median annual values of incremental and breakthrough innovation are one and zero, respectively (Table 3). Vuong (1989) tests confirm the appropriateness of that model for our data. Not surprisingly, we find that the previous year's incremental innovation output is a significant determinant of incremental innovation in the current year ($\beta_{41} = .02$, $p < .01$). Firm size ($\beta_{42} = .30$, $p < .01$) and fixed assets intensity ($\beta_{44} = -.72$, $p < .01$) are also significant determinants of incremental innovation.

For the breakthrough equation, we find that the lagged value of breakthrough innovation ($\beta_{51} = .12$, $p < .01$), firm size ($\beta_{52} = .33$, $p < .01$), and fixed assets intensity ($\beta_{54} = -.61$, $p < .01$) are all significant determinants of current breakthrough levels. Notably, organizational slack is a significant determinant of breakthrough ($\beta_{53} = 1.54$, $p < .05$) but not of incremental innovation. A possible explanation is that incremental innovation is more likely to be incorporated into a firm's baseline investment strategy and less likely to depend on uncertain slack resources. In contrast, breakthrough innovation may be viewed as a more opportunistic activity to be undertaken only when greater slack resources are available.

The results from Equations 4 and 5 indicate that size and resources are important determinants of both breakthrough and incremental innovations. In line with previous research (e.g., Chandy and Tellis 2000; Sorescu, Chandy, and Prabhu 2003), this suggests that large firms introduce both more breakthrough and more incremental innovations. This may explain the large correlation between these two types of innovation ($\rho = .71$, $p < .01$).

Robustness Tests and Additional Analysis

Multidimensional nature of breakthrough innovation. We verify that our results are robust to alternative definitions of breakthrough innovation. Recall that Productscan uses six different breakthrough innovativeness dimensions. In our main analysis, a product is labeled as a breakthrough if it is innovative on one or more of these six dimensions, and each breakthrough carries equal weight, regardless of its number of innovative dimensions. Because 358 of the 1567 breakthroughs are innovative on more than one dimension, we experiment with an alternative approach in which we assign each breakthrough a weight that is proportional to its number of dimensions. The magnitude of the coefficients of interest and their significance levels remain substantively unchanged throughout Equations 1–5.

We also experiment with using the separate innovativeness dimensions that Productscan provided. We reestimate Equations 1–5 six different times, each time replacing the overall breakthrough variable with a narrower variable that is defined exclusively along one of the six dimensions: formulation, merchandising, new market, packaging, positioning, and technology. We find that formulation, packaging, and positioning innovations are positively and significantly ($p < .05$) associated with both Tobin's q and abnormal returns. We also find that each dimension has a positive effect on risk, with the exception of technological innova-

TABLE 5
Innovation, Performance, and Risk

A: Determinants of Normal Profits, Economic Rents, and Risk

Dependent Variables	Independent Variables: Breakthrough Innovation$_{i,t}$	(Breakthrough Innovation)$^2_{i,t}$	Incremental Innovation$_{i,t}$	(Incremental Innovation)$^2_{i,t}$	(Breakthrough × Incremental)$_{i,t}$	Advertising$_{i,t}$	Firm size$_{i,t}$	ROA$_{i,t}$	Model Statistics: Log(Likelihood)		Wald χ^2
Tobin's $q_{i,t}$	.0337**		.000751***		2.69×10^{-6}	.000464*	−.0213		−1097.78		134.0***
(N = 981)	(.0137)		(.000118)		(2.98×10^{-6})	(.000261)	(.0270)				
	.1304***	−.00275***	.001513***	-2.03×10^{-7}***	-2.26×10^{-7}	.000544**	−.1093***		−1082.43		141.6***
	(.0290)	(.000730)	(.000374)	(7.57×10^{-8})	(3.28×10^{-6})	(.000261)	(.0299)				
									R^2 Within	R^2 Between	Wald χ^2
BHAR$_{i,t}$	.000847**		−.0000151		3.71×10^{-6}	.000770		.0000342	.1005	.0753	119.2***
(N = 906)	(.000365)		(.0000575)		(5.11×10^{-6})	(.00438)		(.000338)			
	.00217***	−.0000448***	−.0000179	1.32×10^{-7}	1.23×10^{-6}	.000657		.0000363	.1005	.0886	135.2***
	(.000572)	(.0000136)	(.0000841)	(3.13×10^{-7})	(5.42×10^{-6})	(.00436)		(.000338)			
									Log(Likelihood)		Wald χ^2
Risk$_{i,t}$	.000745***		1.18×10^{-6}		-1.24×10^{-8}	-3.85×10^{-6}	−.00580***	−.000256***	3285.24		1116.1***
(N = 1011)	(.00160)		(1.48×10^{-6})		(3.75×10^{-8})	(2.63×10^{-6})	(.000253)	(.0000327)			
	.00117***	−.0000225***	4.25×10^{-6}	-7.74×10^{-10}	-1.74×10^{-8}	-2.69×10^{6}	−.00612***	−.000290***	3293.90		1162.7***
	(.000368)	(8.05×10^{-6})	(4.93×10^{-6})	(1.002×10^{-11})	(4.34×10^{-8})	(2.72×10^{-6})	.000261	(.0000338)			

B: Summary of Results

Type of Innovation	Empirical Relationship with: Normal Profits	Economic Rents	Firm Risk
Incremental	Positive	Null	Null
Breakthrough	Positive	Positive	Positive

*$p < .10$.
**$p < .05$.
***$p < .01$.
Notes. Standard errors are in parentheses below each coefficient The model includes year dummies, some of which are significant at the $p < .01$ level.

TABLE 6
Firm-Level Breakthrough and Incremental Innovation

A: Determinants of Innovation

Dependent Variables	Independent Variables				
Incremental innovation$_{i,t}$	Incremental innovation$_{i,t-1}$	Firm size$_{i,t-1}$	Organizational slack$_{i,t-1}$	Fixed assets intensity$_{i,t-1}$	Acquisitions$_{i,t-1}$
	.02 (.001)**	.30 (.02)**	.73 (.43)	−.72 (.10)**	.03 (.02)
Model statistics: log-likelihood = −3314.88; χ^2 = 1167.8, $p < .01$					
Breakthrough innovation$_{i,t}$	Breakthrough innovation$_{i,t-1}$	Firm size$_{i,t-1}$	Organizational slack$_{i,t-1}$	Fixed assets intensity$_{i,t-1}$	Acquisitions$_{i,t-1}$
	.12 (.02)**	.33 (.03)**	1.54 (.74)*	−.61 (.20)**	.007 (.03)
Model statistics: log-likelihood = −1212.62; χ^2 = 334.18, $p < .01$					

B: Changes in Innovation Output Through Time

	Breakthrough Innovation	Incremental Innovation
Number of firms with persistent above-average innovation output throughout 1985–2003	1[a]	6
Number of firms with persistent below-average innovation output throughout 1985–2003	93	102

*$p < .05$.
**$p < .01$.
[a] Procter & Gamble.
Notes: Standard errors are in parentheses next to each coefficient.

tions.[14] Overall, we conclude that our results are robust to alternative operationalizations of the breakthrough innovation variable.

Investigating innovative output relative to competitors. We examine the persistence of each firm's innovativeness output relative to that of other firms in the sample; the results appear in Panel B of Table 6. We find that few firms can persistently produce a substantial flow of either breakthrough or incremental innovations. For example, only one firm, Procter & Gamble, had an above-average output of breakthrough innovations each year during our sample period. To ensure that Procter & Gamble does not exert an undue influence on our empirical results, we repeat our entire analysis on a subsample that excludes this firm. The results remain essentially unchanged, suggesting that they are not driven by the sample's top innovator.

Additional analyses using alternative measures of BHARs. Instead of using BHARs adjusted for size, book-to-market, and momentum, we repeat our analysis with a set of industry-adjusted BHARs (Barber and Lyon 1997) and a set of BHARs adjusted only for size. The results remain substantively unchanged.

Decomposing risk into systematic and idiosyncratic components. The seminal work in portfolio theory by Markovitz (1952), Treynor (1962), Sharpe (1964), and Lintner (1965) has decomposed total risk into systematic and idiosyncratic components, which have been studied separately because of their differing effects on various corporate stakeholders. The first component, systematic risk, typically affects only the firm's shareholders. Portfolio theory argues that investors are able to hold diversified portfolios, which eliminates part (but not all) of total firm risk. The more securities held in a portfolio, the lower is the standard deviation of the portfolio's return (Markovitz 1952). However, there is a limit. Even if investors were to hold the ultimate diversified portfolio comprising all publicly traded stocks (i.e., the market portfolio), there remains some residual volatility. This is the risk of the stock market as a whole, compared with the safer bond market. Systematic risk measures the extent to which an individual stock contributes to the risk of this market portfolio. The amount of risk that remains after systematic risk is accounted for is idiosyncratic risk, the second component of total risk. Intuitively, it corresponds to the volatility that can be diversified away by investors who hold large portfolios.

As noted previously, in our main analysis, we use total risk as a dependent variable in Equation 3 because this measure fully captures the effect of innovation on the entire organization. However, in line with recent research in marketing that examines systematic risk (e.g., Madden, Fehle, and Fournier 2006; McAlister, Srinivasan, and Kim 2007), we reestimate Equation 3 using the systematic and idiosyncratic risk components as dependent variables. We compute the two components through annual regressions of firm-specific stock returns on the returns of the market portfolio (not shown). The coefficient of the market portfolio, β, corresponds to the systematic risk component (McAlister, Srinivasan, and Kim 2007). The residual from this same regression is the idiosyncratic component (Ang et al. 2006).

---

[14] Estimating a model that includes all dimensions simultaneously would lead to extreme multicollinearity. In addition, the absence of statistical significance for the "merchandising," "new market," and "technology" innovativeness categories may well reflect the lack of statistical power rather than the absence of a real economic effect. Indeed, none of these three categories has more than 50 observations.

We find that neither breakthrough nor incremental innovations are significantly associated with systematic risk. This suggests that the positive abnormal returns uncovered in Equation 2 are the result of higher cash flows, not lower discount rates, consistent with our economic rents hypothesis. In contrast, the results for idiosyncratic risk are almost identical to those we obtained using total risk. Note that idiosyncratic risk is positively related to breakthrough innovation but unrelated to incremental innovation, suggesting that our total risk results are driven by the idiosyncratic component. This is intuitive because the probability of success of an individual breakthrough innovation is firm specific and likely to be unrelated to macroeconomic risk factors.

Summary and Discussion

To our knowledge, this research is the first to uncover important differences between the role of breakthrough and incremental innovation in building firm value, using three complementary facets of firm performance. We find that incremental innovation is associated with increases in firms' normal profits but does not affect risk and is not a source of economic rents. In contrast, breakthrough innovation is associated with increases in economic rents as well as in risk levels. However, at least from the shareholders' perspective, the increase in risk is compensated by the increase in equity value.

Implications for Theory

Several researchers have decried the scarcity of marketing investments mentioned in boardroom discussions and have called for more research using CEO-relevant metrics (Rust et al. 2004; Srivastava and Reibstein 2004). We subscribe to this stream of research and provide an in-depth examination of the effect of innovation on not just one, but three complementary performance metrics that are relevant to corporate officers. Prior studies that use single metrics of firm performance and limited, convenience samples of innovations have sometimes reached contradictory conclusions on the effect of innovation on firm performance. Using a novel methodological approach and a comprehensive innovation database, we show that incremental innovations help preserve firm value, but breakthroughs are critical for building firm value.

From an industrial organization perspective, the absence of economic rents associated with incremental innovations suggests that firms are unable to attain a strong competitive advantage by launching incremental new products. In general, such products are easily imitable and, thus, competitively priced. In contrast, the positive rents associated with breakthrough innovations suggest that these products enjoy a de facto monopoly protection, perhaps through patents or simply through superior product differentiation, which affords the potential for premium pricing.

Our findings also have implications for the corporate governance literature. Previous research on the relationship between governance and innovation has focused mostly on the size of research and development investment (e.g., Baysinger, Kosnik, and Turk 1991; Hoskisson and Hitt 1988) or the locus of innovation, internal or external (e.g., Hoskisson, Johnson, and Grossman 2002). The current study provides a new addition to the literature. We uncover a significant trade-off that managers must make in the pursuit of breakthrough innovation. On the one hand, breakthroughs are rewarding for shareholders, even when the associated increase in firm risk is taken into account. On the other hand, this additional risk may negatively affect the welfare of other stakeholders, who might conceivably prefer lower levels of breakthrough innovation. Ultimately, the propensity to innovate may depend on how well management incentives are aligned with those of shareholders. We believe that firms with shareholder-oriented governance structures are more likely to foster an organizational culture conducive to breakthrough innovation. Further research could explore the link between the propensity to innovate and various governance variables, such as managerial compensation, board composition, antitakeover provisions, and equity ownership.

Finally, an important implication of our study is that marketing innovation is just as worthy of attention in the boardroom as technological innovation. We find that packaging and merchandising innovations, even though they are not technological marvels, can be the source of significant economic rents. A case in point is Pampryl, which was Pernod Ricard's fruit juice division until 2001. Pampryl achieved a major turnaround in its declining fortunes by introducing a clear plastic packaging technology that prevented oxidation of pasteurized juice. This breakthrough packaging innovation catapulted Pampryl from a brand that was struggling for survival to second position in France's highly competitive juice and nectar market (Kashani 2000). Thus, our results highlight the need for theory development and empirical research on the antecedents and consequences of nontechnological innovations.

Additional research directions stem from our findings. There is undoubtedly a difference in quality between the various products covered by Productscan, and this difference could add additional explanatory power to our three measures of firm performance (Rust, Zahorik, and Keiningham 1995). For example, it would be worthwhile to know whether quality can substitute for scale in product innovation and whether firms can earn economic rents by launching fewer products, but of higher quality.

Our research also implies the possibility that, at least in the CPG sector, a good portion of a brand's equity lies in the innovations that sustain the brand and extend it through time. Strong, financially valuable brands could well be those associated with innovative, flagship products that are perceived as unique and high quality, and this perception may spill over into product line extensions. Further research might examine how branding decisions accompany the market deployment of innovations and ultimately affect their financial value.

Implications for Practice

Our study has clear implications for both corporate boards and management. To members of the board, we note that breakthrough innovation pays off, but at the cost of an increase in firm risk. The additional risk increases the likelihood of financial distress, with potentially negative repercussions on managers, employees, and other firm stakeholders. Although boards want managers to act in shareholders' interests, managers may have a natural tendency to prioritize their own interests and those of other firm's stakeholders, given the disproportionate amount of time they spend interacting with employees, customers, and suppliers compared with shareholders (Grinblatt and Titman 1998). If management is directly concerned with the utility of these other stakeholders, it may be reluctant to pursue breakthrough innovation to the detriment of firm value and shareholder wealth. To preempt this underinvestment problem, boards should align the incentives of management with those of shareholders, perhaps through incentive-compatible executive compensation packages that include equity and stock options.

To CEOs and other corporate officers, we note that incremental innovation keeps firms in business, but breakthrough innovation is the key to achieving sustained long-term growth. Our model allows managers to quantify the average NPV of a breakthrough innovation. Specifically, the estimated coefficient β_{22} in Equation 2 is .000847, which means that, on average, each breakthrough innovation is associated with a .0847% increase in the market value of the firm's equity. This increase accounts for all costs (including a normal equity charge) and thus represents pure economic profit (or rents).[15] For the typical firm in our sample with an average market value of equity of $4.9 billion, the magnitude of β_{22} implies that the NPV of economic rents associated with a single breakthrough innovation is approximately $4.2 million.

A word of caution, however, is necessary. Our finding of positive rents for breakthrough innovations seems to suggest that there is no upper limit on the amount of breakthroughs a firm should undertake. However, such inference would be misleading. Our method implicitly assumes that firms are able to provide in-house financing for breakthrough innovations. In other words, we assume that funds spent on research and development come primarily from cash that is available within the firm rather than from external financing. The significant, positive coefficient on the slack variable in Equation 5 is consistent with this assumption and suggests that in-house fund availability is indeed a factor in determining the level of breakthrough activity. Raising external capital is always costly (Grinblatt and Titman 1998), and firms that exhaust all in-house funding resources must carefully weigh the additional benefits of breakthrough innovation against the fixed costs of external financing. Thus, a more realistic interpretation of our results is that managers should pursue breakthrough innovation to the point at which they can finance these costs internally. Whether the benefits of breakthrough innovation are sufficiently high to offset the costs of external financing is a topic for further research.

Appendix
Instrumental Variables Estimation

We describe our instrumental variables estimation procedure for the linear versions of Equations 1 and 3. The procedure is identical for the nonlinear versions, and in the interest of brevity, we do not repeat it here.

Let I* and B* be the predicted values of Incremental $\text{innovation}_{i,t}$ and Breakthrough $\text{innovation}_{i,t}$ obtained from Equations 4 and 5, respectively. Let IB* be the predicted value of the product between Incremental $\text{innovation}_{i,t}$ and Breakthrough $\text{innovation}_{i,t}$ obtained from Equation A1 (see Kelejian 1971):

$$
\begin{aligned}
\text{(A1)}\quad & (\text{Incremental innovation} \times \text{Breakthrough innovation})_{i,t} \\
& = \alpha_{60} + \beta_{61}(\text{Incremental innovation} \times \text{Breakthrough} \\
& \quad \text{innovation})_{i,t-1} \\
& + \beta_{62}\text{Firm size}_{i,t-1} + \beta_{63}\text{Organizational slack}_{i,t-1} \\
& + \beta_{64}\text{Fixed asset intensity}_{i,t-1} + \beta_{65}\text{Acquisitions}_{i,t-1} \\
& + \varepsilon_{6i,t}.
\end{aligned}
$$

We use I*, B*, and IB* directly in Equations 1 and 3 in place of the actual variables. This produces Equations 1* and 3*.

$$
\begin{aligned}
(1^*)\quad & \text{Tobin's } q_{i,t} = \alpha^*_{10} + \beta^*_{11}I^*_{i,t} + \beta^*_{12}B^*_{i,t} + \beta^*_{13}IB^*_{i,t} \\
& + \beta^*_{14}\text{Advertising}_{i,t} + \beta^*_{15}\text{Firm size}_{i,t} + \varepsilon^*_{1i,t}, \text{ and}
\end{aligned}
$$

$$
\begin{aligned}
(3^*)\quad & \text{Risk}_{i,t} = \alpha^*_{30} + \beta^*_{31}I^*_{i,t} + \beta^*_{32}B^*_{i,t} + \beta^*_{33}IB^*_{i,t} \\
& + \beta^*_{34}\text{Advertising}_{i,t} + \beta^*_{35}\text{ROA}_{i,t} + \varepsilon^*_{3i,t}.
\end{aligned}
$$

Greene (2003, pp. 397–400) and Maddala (1992, pp. 373–74) explain that coefficients β^* obtained in this manner are consistent, but their standard errors (s*) are biased toward zero. Thus, relying on s* for inferences could overreject the null hypothesis. However, Maddala (1992, p. 376) shows that standard errors (s*) can be corrected by multiplying them with an inflation factor ψ, which is unique to each equation. We describe the procedure for computing the inflation factor for Equation 1*; a similar procedure applies to Equation 3* and to the nonlinear versions of Equations 1* and 3*.

Following Maddala (1992), we compute the inflation factor as

$$\psi = \sigma_u/\sigma_w,$$

---

[15] The term "equity charge" refers to the opportunity cost of equity—that is, the rate of return that stock market investors require times the dollar value of the investment.

where σ_w is the standard deviation of residuals ($\varepsilon^*_{1i,t}$) from Equation 1* and σ_u is the standard deviation of a pseudoresidual ($\varepsilon^{**}_{1i,t}$), given by

$$
\begin{aligned}
\text{(A2)}\quad \varepsilon^{**}_{1i,t} = \text{Tobin's } q_{i,t} - \{&\alpha^*_{10} \\
&+ \beta^*_{11}\text{Incremental innovation}_{i,t} \\
&+ \beta^*_{12}\text{Breakthrough innovation}_{i,t} \\
&+ \beta^*_{13}(\text{Incremental innovation}_{i,t} \\
&\times \text{Breakthrough innovation}_{i,t}) \\
&+ \beta^*_{14}\text{Advertising}_{i,t} + \beta^*_{15}\text{Firm size}_{i,t}\}.
\end{aligned}
$$

That is, the pseudoresiduals computed in Equation A2 use the estimated coefficients from Equation 1* combined with the actual innovation variables (not the instrumented ones). In our empirical estimation, we found that the inflation factors were low, ranging from 1.0086 to 1.0125 (an inflation factor equal to unity implies that no adjustment is needed).

REFERENCES

Aaker, David and Robert Jacobson (1994), "The Financial Information Content of Perceived Quality," *Journal of Marketing Research*, 31 (May), 191–202.

Advertising Age (2005), "Does Innovation Pay?" (August 15), 4.

Anderson, Philip and Michael L. Tushman (1990), "Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change," *Administrative Science Quarterly*, 35 (4), 604–633.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang (2006), "The Cross-Section of Volatility and Expected Returns," *Journal of Finance*, 61 (1), 259–99.

Barber, Brad and John Lyon (1997), "Detecting Long-Run Abnormal Stock Returns: The Empirical Power and Specification of Test Statistics," *Journal of Financial Economics*, 43 (3), 341–72

Baysinger, Barry D., Rita D. Kosnik, and Thomas A. Turk (1991), "Effects of Board Ownership Structure on Corporate R&D Strategy," *Academy of Management Journal*, 34 (1), 205–214.

Bayus, Barry L., Gary Erickson, and Robert Jacobson (2003), "The Financial Rewards of New Product Introductions in the Personal Computer Industry," *Management Science*, 49 (2), 197–210.

Beck, Nathaniel and Jonathan N. Katz (1995), "What to Do (and Not to Do) with Time-Series Cross-Section Data," *American Political Science Review*, 89 (3), 643–47.

Blundell, Richard, Rachel Griffith, and John Van Reenen (1999), "Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms," *Review of Economic Studies*, 66 (3), 529–54.

Boston Consulting Group (2005), *Innovation 2005*. Boston: Boston Consulting Group.

Bureau of Economic Analysis (2007), "National Income and Product Accounts Table 2.3.5: Personal Consumption Expenditures by Major Type of Product," (accessed February 2, 2007), [available at http://bea.gov/bea/dn/nipaweb/].

Chandy, Rajesh K. and Gerard J. Tellis (1998), "Organizing for Radical Product Innovation: The Overlooked Role of Willingness to Cannibalize," *Journal of Marketing Research*, 35 (November), 474–87.

——— and ——— (2000), "The Incumbent's Curse? Incumbency, Size, and Radical Product Innovation," *Journal of Marketing*, 64 (July), 1–17.

Chaney, Paul K., Timothy M. Devinney, and Russell S. Winer (1991), "The Impact of New Product Introductions on the Market Value of Firms," *Journal of Business*, 64 (4), 573–610.

Christensen, Clayton M. (1997), *The Innovator's Dilemma*. New York: Harper Business.

——— and Joseph L. Bower (1996), "Customer Power, Strategic Investment, and the Failure of Leading Firms," *Strategic Management Journal*, 17 (3), 197–218.

Cooper, Robert G. (2005), "Your NPD Portfolio May Be Harmful to Your Business Health," *PDMA Visions*, 29 (2), 22–26.

Copeland, Thomas E. and J. Fred Weston (1992), *Financial Theory and Corporate Policy*, 3d ed. Reading, MA: Addison Wesley.

Cyert, Richard M. and James G. March (1963), *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice Hall.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers (1997), "Measuring Mutual Fund Performance with Characteristic-Based Benchmarks," *Journal of Finance*, 52 (3), 1035–1058.

Davis, Gerald F. and Suzanne K. Stout (1992), "Organization Theory and the Market for Corporate Control: A Dynamic Analysis of the Characteristics of Large Takeover Targets, 1980-1990," *Administrative Science Quarterly*, 37 (4), 605–633.

Dewar, Robert D. and Jane E. Dutton (1986), "The Adoption of Radical and Incremental Innovations. An Empirical Analysis," *Management Science*, 32 (11), 1422–33.

Dixit, Avinash (1980), "The Role of Investment in Entry-Deterrence," *Economic Journal*, 90 (357), 95–106.

Eddy, Albert R. and George B. Saunders (1980), "New Product Announcements and Stock Prices," *Decision Sciences*, 11 (1), 90–97.

Fama, Eugene and Kenneth R. French (1992), "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47 (2), 427–65.

Foster, Richard N. and Sarah Kaplan (2001), *Creative Destruction: Why Companies That Are Built to Last Underperform the Market—and How to Successfully Transform Them*. New York: Currency.

Garcia, Rosanna and Roger Calantone (2002), "A Critical Look at Technological Innovation Typology and Innovativeness Terminology: A Literature Review," *Journal of Product Innovation Management*, 19 (2), 110–32.

Gatignon, Hubert and Thomas S. Robertson (1985), "A Propositional Inventory for New Diffusion Research," *Journal of Consumer Research*, 11 (4), 849–67.

Geroski, Paul, Steve Machin, and John Van Reenen (1993), "The Profitability of Innovating Firms," *RAND Journal of Economics*, 24 (2), 198–211.

Golden, Brian R. (1992), "The Past Is the Past: Or Is It? The Use of Retrospective Accounts as Indicators of Past Strategy," *Academy of Management Journal*, 35 (4), 848–60.

Gompers, Paul, Joy Ishii, and Andrew Metrick (2003), "Corporate Governance and Equity Prices," *Quarterly Journal of Economics*, 118 (1), 107–155.

Gourville, John T. (2005), "The Curse of Innovation: Why Innovative New Products Fail," Marketing Science Institute Report No. 05-117.

Greene, William H. (2003), *Econometric Analysis*, 5th ed. Upper Saddle River, NJ: Prentice Hall.

Grinblatt, Mark and Sheridan Titman (1998), *Financial Markets & Corporate Strategy*. Homewood, IL: Richard D. Irwin.

Hambrick, Donald C. and Ian C. McMillan (1985), "Efficiency of Product R&D in Business Units: The Role of Strategic Context," *Academy of Management Journal*, 28 (3), 527–47.

Henard, David H. and David M. Szymanski (2001), "Why Some New Products Are More Successful Than Others," *Journal of Marketing*, 38 (July), 362–75.

Herold, David M., Narayanan Jayaraman, and C.R. Narayanaswamy (2003), "What Is the Relationship Between Organizational Slack and Innovation?" *Journal of Managerial Issues*, 18 (3), 372–92.

Hill, Charles W.L. and Frank T. Rothaermel (2003), "The Performance of Incumbent Firms in the Face of Radical Technological Innovation," *Academy of Management Review*, 28 (2), 257–74.

Hirshleifer, David and Y. Suh (1992), "Risk, Managerial Effort and Project Choice," *Journal of Financial Intermediation*, 2 (September), 308–345.

Hoskisson, Robert E. and Michael A. Hitt (1998), "Strategic Control Systems and Relative R&D Investment in Large Multi-Product Firms," *Strategic Management Journal*, 9 (6), 605–621.

———, ———, Richard A. Johnson, and Wayne Grossman (2002), "Conflicting Voices: The Effects of Institutional Ownership Heterogeneity and Internal Governance on Corporate Innovation Strategies," *Academy of Management Journal*, 45 (4), 697–716.

Jegadeesh, Narasimhan and Sheridan Titman (1993), "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48 (1), 65–91.

Jensen, Michael and William Meckling (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," *Journal of Financial Economics*, 3 (4), 305–360.

Joshi, Amit and Dominique M. Hanssens (2004), "Advertising Spending and Market Capitalization," Marketing Science Institute Report No. 04-110.

Kaplan, Steven N. and Luigi Zingales (1997), "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?" *Quarterly Journal of Economics*, 112 (1), 169–215.

Kashani, Kamran (2000), "Brand New Growth: Innovating Your Brand to Grow," *IMD Newsletter*, 74 (December), 1–4.

Kelejian, Harry H. (1971), "Two-Stage Least Squares in Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables," *Journal of the American Statistical Association*, 66 (334), 373–74.

Kennedy, Peter (2003), *A Guide to Econometrics*. Cambridge, MA: The MIT Press.

Kim, W. Chan and Renée Mauborgne (1997), "Value Innovation: The Strategic Logic of High Growth," *Harvard Business Review*, 75 (1), 103–112.

Lee, Hun, Ken G. Smith, Curtis M. Grimm, and August Schomburg (2000), "Timing, Order and Durability of New Product Advantages with Imitation," *Strategic Management Journal*, 21 (1), 23–30.

Lintner, John (1965), "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, 47 (1), 13–37.

Long, Scott J. and Jeremy Freese (2003), *Regression Models for Categorical Dependent Variables Using Stata*. College Station, TX: Stata Press.

Maddala, G.S. (1992), *Introduction to Econometrics*, 2d ed. New York: Macmillian.

Madden, Thomas J., Frank Fehle, and Susan Fournier (2006), "Brands Matter: An Empirical Demonstration of the Creation of Shareholder Value Through Branding," *Journal of the Academy of Marketing Science*, 34 (2), 224–35.

Markowitz, Harry M. (1952), "Portfolio Selection," *Journal of Finance*, 7 (1), 77–91.

Marshall, Alfred (1920), *Principles of Economics*. London: Macmillan.

Maximovic, Vojislav and Sheridan Titman (1991), "Financial Policy and Reputation for Product Quality," *Review of Financial Studies*, 4 (1), 175–200.

McAlister, Leigh, Raji Srinivasan, and MinChung Kim (2007), "Advertising, Research and Development, and Systematic Risk of the Firm," *Journal of Marketing*, 71 (January), 35–48.

Miles, Raymond E. and Charles C. Snow (1978), *Organizational Strategy, Structure and Process*. New York: McGraw-Hill.

Min, Sungwook, Manohar U. Kalwani, and William T. Robinson (2006), "Market Pioneer and Early Follower Survival Risks: A Contingency Analysis of Really New Versus Incrementally New Product-Markets," *Journal of Marketing*, 70 (January), 15–33.

Neff, Jack (2005), "Small Ball: Marketers Rely on Line Extensions," *Advertising Age*, (April 11), 10–12.

Pauwels, Koen, Jorge Silva-Risso, Shuba Srinivasan, and Dominique M. Hanssens (2004), "New Products, Sales Promotions, and Firm Value: The Case of the Automobile Industry," *Journal of Marketing*, 68 (October), 142–56.

Ronn, Ehud I. and Avinash K. Verma (1986), "Pricing Risk Adjusted Deposit Insurance: An Option-Based Model," *Journal of Finance*, 41 (4), 871–95.

Rust, Roland T., Tim Ambler, C.S. Carpenter, V. Kumar, and Rajendra K. Srivastava (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (October), 76–89.

———, Anthony J. Zahorik, and Timothy L. Keiningham (1995), "Return on Quality (ROQ): Making Service Quality Financially Accountable," *Journal of Marketing*, 59 (April), 58–70.

Schrage, Michael (2000), "Getting Beyond the Innovation Fetish," *Fortune*, (November), 225–32.

Schwartz, Jonathan (2006), "The Five Founding Principles That Drive Innovation," *The Financial Times*, (September 12), (accessed January 15, 2007), [available at http://www.FT.com].

Sharma, Anurag and Lacey Nelson (2004), "Linking Product Development Outcomes to Market Valuation of the Firm: The Case of the U.S. Pharmaceutical Industry," *Journal of the Academy of Marketing Science*, 21 (5), 297–308.

Sharpe, William F. (1964), "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance*, 19 (3), 425–42.

Singh, Manohar, Sheri Faircloth, and Ali Nejadmalayeri (2005), "Capital Market Impact of Product Marketing Strategy: Evidence from the Relationship Between Advertising Expenses and Cost of Capital," *Journal of the Academy of Marketing Science*, 33 (4), 432–44.

Sorescu, Alina B., Rajesh K. Chandy, and Jaideep C. Prabhu (2003), "Sources and Financial Consequences of Radical Innovation: Insights from Pharmaceuticals," *Journal of Marketing*, 67 (October), 82–101.

———, ———, and ——— (2007), "Why Some Acquisitions Do Better Than Others: Product Capital as a Driver of Long-Term Stock Returns," *Journal of Marketing Research*, 44 (February), 57–72.

———, Venkatesh Shankar, and Tarun Kushwaha (2007), "New Product Preannouncements and Shareholder Value: Don't Make Promises You Can't Keep," *Journal of Marketing Research*, 44 (August), 468–89.

Spence, A. Michael (1977), "Entry, Capacity, Investment and Oligopolistic Pricing," *Bell Journal of Economics*, 8 (2), 534–44.

Srinivasan, Shuba, Koen Pauwels, Jorge Silva-Risso, and Dominique M. Hanssens (2006), "Product Innovations, Advertising Spending and Stock Returns," Marketing Science Institute Report No. 06-110.

Srivastava, Rajendra and David J. Reibstein (2004), "Metrics for Linking Marketing to Financial Performance," Marketing Science Institute Report No. 05-200e.

Tirole, Jean (1988), *The Theory of Industrial Organization*. Cambridge, MA: The MIT Press.

Titman, Sheridan (1984), "The Effect of Capital Structure on the Firm's Liquidation Decision," *Journal of Financial Economics*, 13 (1), 137–52.

Treacy, Michael (2004), "Innovation as a Last Resort," *Harvard Business Review*, 82 (7–8), 29–30.

Treynor, Jack L. (1962), "Toward a Theory of Market Value of Risky Assets," unpublished manuscript. (A final version was published in 1999, in *Asset Pricing and Portfolio Performance: Models, Strategy and Performance Metrics*, Robert A. Korajczyk, ed. London: Risk Books, 15–22.)

VanderWerf, Pieter A. and John F. Mahon (1997), "Meta-Analysis of the Impact of Research Methods on Findings of First-Mover Advantage," *Management Science*, 43 (11), 1510–19.

Von Neumann, John and Oskar Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Vuong, Quang H. (1989), "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57 (2), 307–333.

Wermers, Russ (2003), "Is Money Really 'Smart'? New Evidence on the Relation Between Mutual Fund Flows, Manager Behavior and Performance Persistence," working paper, Robert H. Smith School of Business, University of Maryland, (May), (accessed June 1, 2007), [available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=414420#].

Weston, J. Fred (1950), "A Generalized Uncertainty Theory of Profit," *American Economic Review*, 40 (1), 40–60.

White, Halbert (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 48 (4), 817–38.

Zhou, Kevin Zheng, Chi Kin Yim, and David K. Tse (2005), "The Effects of Strategic Orientations on Technology- and Market-Based Breakthrough Innovations," *Journal of Marketing*, 69 (2), 42–60.

Christian Homburg, Ove Jensen, & Harley Krohmer

Configurations of Marketing and Sales: A Taxonomy

Little is known about the interface between separate marketing units and sales units. This article develops a multidimensional model of the marketing and sales interface. The model integrates a broad range of conceptual domains, including information sharing, structural linkages, power, orientations, and knowledge of marketing and sales. The authors empirically explore the conceptual model through a cross-industry study of 337 European Union–based companies. They identify five empirical archetypes of the marketing and sales interface. The taxonomy shows that the role and characteristics of marketing and sales vary a great deal. This finding challenges existing stereotypes about marketing and sales. Finally, the article explores organizational outcomes of the five configurations. The findings suggest that the most successful configurations are characterized by strong structural linkages between marketing and sales and a high extent of market knowledge in marketing.

Keywords: marketing organization, sales organization, interface, taxonomy, configuration

The interface between marketing units and sales units is receiving increasing conceptual attention (Dewsnap and Jobber 2000, 2002; Rouziès et al. 2005). In the past, empirical studies have typically not distinguished between marketing and sales (hereinafter M & S) but have subsumed both under the term "marketing organization" (Dastmalchian and Boag 1990; Ruekert and Walker 1987a, b). This contrasts sharply with managerial reports that "as a rule,… they're separate functions within an organization" (Kotler, Rackham, and Krishnaswamy 2006, p. 68). Workman, Homburg, and Gruner (1998, p. 37) state that "more than 30 years after the call to integrate sales and marketing activities,… we find no firms that had adopted this recommendation.… Additional research is needed to explore further the relationship between [M & S]."

According to anecdotal accounts, the separation between M & S runs far deeper than reporting relationships and organizational charts: M & S are also separated by different orientations and knowledge (Deshpandé and Webster 1989). It has been claimed that marketing is product focused and long-term focused, whereas sales is customer focused and short-term focused (Cespedes 1995; Rouziès et al. 2005). It has been alleged that marketing lives in an "ivory tower," whereas sales has been lauded for its market knowledge (Carpenter 1992; Donath 2004). Overall, the M & S relationship is reported to be far from harmonious (Montgomery and Webster 1997).

However, caution is warranted when general statements are made about M & S. The literature indicates that the interface between M & S varies a great deal across companies and industries. Whereas the orientations and knowledge of M & S are far apart in some companies, they are aligned in others (Carpenter 1992; Lorge 1999). Whereas sales is centered on products in some firms, it is centered on customer groups in others (Homburg, Workman, and Jensen 2000; Webster 1997). Whereas marketing has a driving role in some consumer goods firms, it has a limited role in some high-tech firms (Cespedes 1993; Workman 1993). Research by Homburg, Workman, and Krohmer (1999), Piercy (1986), and Starr and Bloom (1994) confirms a large degree of variation in the relative influence of M & S.

To our knowledge, this variation of the M & S interface across companies has never been systematically investigated. As our literature review shows, some authors have conceptually discussed typical evolutionary stages of marketing within the organization (Day 1999; Kotler, Rackham, and Krishnaswamy 2006; Webster 1997). However, these discussions have focused on selective issues, such as power, and are not based on clearly defined conceptual dimensions. There has never been a study that empirically classifies the existing varieties of M & S configurations using a broad conceptual model. There is a need for large-scale studies to understand how separate M & S units are configured. Against this background, our article develops an integrative conceptualization of the M & S interface. On the basis of this conceptualization, we conduct an empirical study of 337 business units in seven industry sectors. Our specific research goals are as follows:

- To identify constructs that constitute M & S configurations from the literature,
- To empirically identify varieties of M & S configurations and develop a taxonomy, and
- To explore the outcomes of the different M & S configurations.

Christian Homburg is Professor of Marketing, Chair of the Marketing Department, and Director of the Institute for Market-Oriented Management, University of Mannheim (e-mail: homburg@bwl.uni-mannheim.de). Ove Jensen is Professor of Marketing and Chair of Business-to-Business Marketing, WHU–Otto Beisheim School of Management (e-mail: ove.jensen@whu.edu). Harley Krohmer is Professor of Marketing, Chairman of the Marketing Department, and Director of the Institute of Marketing and Management, University of Berne (e-mail: prof.krohmer@imu.unibe.ch).

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Our taxonomical research approach is consistent with the configurative approach to organizational analysis that studies complex phenomena through overall patterns of multiple variables rather than through selected variables and their bivariate relationships (Ketchen et al. 1997; Meyer, Tsui, and Hinings 1993). Recent research has applied taxonomies to comprehend how firms relate to their markets and their customers (Cannon and Perreault 2001; Coviello et al. 2002), how sellers manage key accounts (Homburg, Workman, and Jensen 2002), or how marketing contributes to the implementation of a business strategy (Slater and Olson 2001).

We organize the rest of this article as follows: We review literature on the M & S interface. Then, we discuss procedures for selecting input variables to taxonomies and identify the relevant domains and constructs of the M & S interface. Afterward, we document the measurement of M & S constructs and describe the taxonomical procedures. In the "Results" section, we present the taxonomy and explore outcomes of the configurations. The article concludes with implications for academic research and managerial practice.

Literature Review

Much empirical work on marketing organization has not distinguished between marketing units and sales units. Traditionally, sales has been conceptualized as a subunit of the marketing department that reports to a marketing executive (Ruekert, Walker, and Roering 1985; Weitz and Anderson 1981). For example, Ruekert and Walker (1987a, b) and Dastmalchian and Boag (1990) speak about marketing as one whole and barely mention the word sales. This view of M & S as one unit is challenged by qualitative reports (Piercy 1986). Workman, Homburg, and Gruner (1998, p. 37) state that "in the 47 firms we studied, we never observed a sales manager reporting to a marketing manager." The separation of M & S has received more attention only recently. Our review includes empirical studies and selected conceptual articles on the M & S interface.

Empirical Studies

Table 1 summarizes the empirical studies on the M & S interface. From these, we draw three conclusions. First, there is a general lack of empirical insight into the M & S interface. Table 1 contains only eight empirical studies, two of which stem from the same data set. Second, of these eight empirical studies, only three (Cespedes 1995; Strahle, Spiro, and Acito 1996; Workman, Homburg, and Gruner 1998) are dedicated to the organization of the M & S interface. The other five studies just touch on the M & S interface as a side issue. Montgomery and Webster (1997) and Panigyrakis and Veloutsou (1999) discuss marketing's interfaces in general and point out the importance of the interface between M & S. Homburg, Workman, and Krohmer (1999) and Krohmer, Homburg, and Workman (2002) focus on marketing's influence over market-related activities. However, their empirical study explicitly distinguishes between the influence of marketing and the influence of sales. What we can gather from their findings is that many of the activities classically labeled as "marketing activities" are dominated by sales rather than by marketing, most notably distribution strategy and pricing. Smith, Gopalakrishna, and Chatterjee (2006) do not study organizational aspects of M & S, but they model sales response to communication policies of M & S units.

The third, and most important, conclusion from Table 1 is that, with one exception, all the empirical studies focus on an "average" or a typical M & S relationship and do not analyze variation in the organization of M & S. As an example, Strahle, Spiro, and Acito (1996), who note that strategies and plans in M & S units are inconsistent and badly communicated, do not distinguish between different configurations of M & S. As another example, Cespedes's (1995) qualitative report describes typical cognitive orientations of marketing, sales, and service; typical task allocations between these three groups; and typical integration mechanisms. Only Workman, Homburg, and Gruner (1998) explicitly address variation in the organization of M & S. However, their typology is based solely on reporting relationships and organization charts. As Webster (1997, p. 55) notes, "future marketing organizations … will not be possible to visualize … with the tools of the standard organization chart." In addition, they focus on large companies with multiple strategic business units (SBUs). Finally, their sample comprises only 47 firms.

To conclude our review of extant empirical work, there seems to be a lack of large-scale empirical studies of alternative M & S configurations that integrate the broad range of issues studied in isolation hitherto: power (Homburg, Workman, and Krohmer 1999), information sharing (Strahle, Spiro, and Acito 1996), structural integration mechanisms (Cespedes 1995), cognitive orientations, and knowledge (Cespedes 1995; Montgomery and Webster 1997).

Conceptual Articles

Two categories of conceptual articles on separate M & S units can be distinguished. The first category revolves around integration devices for M & S. Drawing on studies of marketing's integration with other functional groups (e.g., finance, production, research and development [R&D]), Dewsnap and Jobber (2000) propose antecedents to M & S integration, such as formalization, decentralization, or category teams. In a similar vein, Rouziès and colleagues (2005) develop a framework of M & S integration. They conceptualize structural, procedural, cultural, and people-based integration mechanisms for M & S. They also propose environmental, customer-related, competitor-related, and internal moderators of the integration–performance link. Finally, taking a social psychology perspective, Dewsnap and Jobber (2002) conceptually explore how goal conflict and departmental identities hamper an effective M & S relationship. Again, we find that these articles are a rich source of key constructs. However, none of these articles differentiate between different M & S configurations whose use of the integration mechanisms varies.

The second category of conceptual articles comprises managerial articles on evolutionary stages of marketing

TABLE 1
Empirical Studies of the M & S Interface

Authors	Focus of Study	Empirical Approach	Empirical Basis	Key Insights into M & S Interface	Analysis of Various M & S Configurations
Cespedes (1995)	Integration among product management, sales management, and customer service	Qualitative	200 interviews in the computing, telecommunications services, consumer packaged goods, and medical products industries	Product management and sales management have different windows to the world. Differences pertain to role alignment (products versus accounts), time horizons, and responsibilities. Differences reflect specialization within each group and are necessary to cope with increasingly complex marketing problems.	No
Homburg, Workman, and Krohmer (1999)	Marketing's influence within the firm	Quantitative	514 SBUs in consumer packaged goods, electrical equipment, and mechanical machinery industry	Marketing has the greatest influence over advertising messages, customer satisfaction management, strategic direction of the business unit, and new product development. Sales has the greatest influence over expansions into new geographic markets, distribution strategy, customer service, and pricing.	No
Krohmer, Homburg, and Workman (2002)	Cross-functional dispersion of marketing activities	Quantitative	Same as Homburg, Workman, and Krohmer (1999)	Influence over many marketing acitivities is spread between M & S. In effective SBUs, both M & S are more influential than R&D, manufacturing, and finance.	No
Montgomery and Webster (1997)	Marketing's interfunctional interfaces	Qualitative	47 attendees of an Marketing Science Institute workshop	There is much conflict between M & S. Sales is driven by short-term revenue, and marketing is driven by long-term profit and margin goals.	No
Panigyrakis and Veloutsou (1999)	Interfaces of brand managers	Quantitative	161 brand managers from 48 consumer goods firms	Sales and marketing are the two interfaces on which that brand managers spend the most time.	No
Smith, Gopalakrishna, and Chatterjee (2006)	Timing and budget allocation of marketing communication for generating sales leads, appointments, and closures	Quantitative	1 home improvement retailer, 6068 purchasing events	There is a complex interplay among marketing communication efforts, delays in follow-up, and sales efficiencies. Importance of follow-up timing implies importance of internal collaboration between M & S.	No
Strahle, Spiro, and Acito (1996)	Planning of market strategies, sales objectives, and sales activities	Quantitative	367 sales managers and 25 marketing executives from 25 firms	In general, sales managers do not set sales objectives that are consistent with the strategy specified by a marketing executive for a particular product. Some of the reasons for this gap are miscommunications and volume-goal differences.	No
Workman, Homburg, and Gruner (1998)	Dimensions and determinants of marketing organizations	Qualitative	72 managers in 47 firms	M & S are typically separate departments and are not integrated under a chief marketing executive. In none of the firms did a sales manager report to a marketing manager. There are five typical structural locations of M & S. M & S subunits are found at the corporate level, at the business unit level, and at the local (country) level.	No, but description of various reporting lines

within the organization. To the best of our knowledge, the articles by Day (1999), Kotler, Rackham, and Krishnaswamy (2006), and Webster (1997) are the closest to what may be referred to as a conceptual typology of M & S interfaces. Thus, it is important to review their classifications here. However, note that these classifications are narrative and do not systematically develop archetypes based on clearly defined dimensions. Note also that the key focus of Day (1999) and Webster (1997) is on marketing's role within a multi-business-unit corporation, not on the M & S interface within a single business unit. Table 2 reconstructs key aspects from the three narrative classifications: We organized the three narrative classifications in terms of five key issues: power, structural linkages, orientations, knowledge, and outcomes of M & S cooperation. However, the three texts do not discuss all these issues for all organizational types. Thus, there are blanks in Table 2. An issue that has received the most attention is marketing's power over market-related activities.

This article builds on these classifications in two ways. First, our conceptual model integrates the conceptual issues raised. For example, the power balance between M & S is identified as an important conceptual dimension. Second, our "Discussion" section compares the empirical results of our study with the conceptual types proposed in the previous sources.

The key limitation of these conceptual classifications is the lack of clearly defined dimensions that enable a systematic comparison of the archetypes. The next section

TABLE 2
Conceptual Classifications of Separate M & S Functions

Authors	Interface Type	Power	Structural Linkages	Orientations	Knowledge	Cooperation Outcomes
Day (1999)	Marketing as a functional fiefdom	Marketing is the traditional lead function, but there is a power shift to sales to be more responsive to market needs		Product-oriented marketing		
	Marketing as a subordinate function	Marketing as sales support	Core process teams	Tactical, implementation-focused marketing		
	Marketing as a central guidance function	Marketing is responsible for strategy, information sharing, and coordination	Business planning, market-oriented organizational structures	Market-oriented culture		Market leadership
Kotler, Rackham, and Krishnaswamy (2006)	Marketing as adjunct to sales	Sales dominates over marketing				Positive relationship between M & S
	Marketing as an independent player	Powerful brand managers		Sales is short-term oriented	Doubts about marketing's competence	Disagreements between M & S
Webster (1997)	Marketing as a distinct function	Marketing pursues excellence in all phases of integrated marketing strategy	Functional silos, weak integration with other business functions			Ongoing conflict between M & S, slow response to market environment
	Marketing integrated with other business functions	Marketing responsible for program development	Teams, functional silos broken down	Customer focus		Most successful form

addresses this limitation and develops a conceptual model of the M & S interface.

Conceptual Model

This section develops our conceptual model of the M & S interface. As we stated in the literature review, there is a need for a multidimensional taxonomy. We first explain which guidelines instructed the construction selection. We then proceed to identify the general conceptual domains of interfaces. Finally, we define specific constructs within the domains.

Procedures for Selecting Constructs

The literature is replete with recommendations for the selection of input variables to a classification (for reviews, see Bailey 1994; Dess, Newport, and Rasheed 1993; Rich 1992). In unexplored territory, such as the M & S interface, the literature unequivocally advises choosing variables "in a way that fosters rich description of a sample's characteristics" (Ketchen and Shook 1996, p. 444). Some authors have understood this as a call to include as many variables as possible (McKelvey 1975). However, simulation studies have warned that spurious variables (i.e., variables that do not differentiate between clusters) distort the detection of cluster structures in the data (Punj and Stewart 1983). For this reason, we winnow two types of variables. First, we identify a parsimonious set of theoretical core constructs that serve as input variables to the cluster algorithm. These constructs are grounded in the literature to ensure their relevance. Second, we analyze several descriptive variables that do not enter the cluster procedures but are used to characterize the clusters.

The literature has also discussed how much interdependence between the cluster variables is admissible. Early articles recommend the exclusion of variables that are logically or empirically correlated (Sneath and Sokal 1973). Recent articles contend that unlike other multivariate methods (e.g., regression analysis), most cluster methods do not rely on the assumption that the input variables are uncorrelated (Arabie and Hubert 1994; Milligan 1996). Thus, we accept some conceptual overlap and empirical correlation between the constructs. However, we ensure discriminant validity in measuring these constructs.

To conclude, we attempt to cover the relevant conceptual domains while keeping the number of variables parsimonious. Toward this end, we use a two-step selection procedure. First, we identify the relevant conceptual domains to achieve conceptual breadth. Second, we select theoretical core constructs within each conceptual domain. Constructs are identified as core if their relevance has been repeatedly emphasized in extant literature on M & S, if they are grounded in theoretical literature on interfaces, and if they are applicable across industries. Given the danger of spurious variables, our selection procedure weighs parsimony and theoretical justification of constructs more heavily than a fully exhaustive list of constructs.

Conceptual Domains of the M & S Interface

The systems perspective suggests that two organizational subsystems are linked by multiple relationships (Bunge 1996; Scott 1992). As Figure 1 indicates, we view interfaces to be constituted by five conceptual domains. Each domain contains one or more conceptual dimensions. The conceptual domains are not constructs of a higher order but merely conceptual categories, or conceptual containers, of similar constructs. We emphasize that the order of the domains in Figure 1 does not suggest a hierarchical structure, especially not in the strong sense that one domain is a functional prerequisite of another.[1]

The first conceptual domain is information sharing. It contains the extent of cross-functional intelligence dissemination and knowledge sharing. The importance of information sharing and communication has been emphasized by literature on intraorganizational interfaces (Fisher, Maltz, and Jaworski 1997; Kahn, Reizenstein, and Rentz 2003) and by Cannon and Perreault's (1999) taxonomy of interorganizational interfaces. In addition, information sharing is emphasized by studies on market orientation, organizational learning, and new product development (Marinova 2004; Moorman 1995).

The second conceptual domain encompasses structural linkages. Structural linkages refer to the extent to which formal horizontal platforms or interaction channels are established for interdepartmental activities (Workman, Homburg, and Gruner 1998). In an intraorganizational context, structural linkages are often presented as a continuum of coordination mechanisms (Martinez and Jarillo 1989). Structural linkages are also an important element in a taxonomy of interorganizational interfaces (Cannon and Perreault 1999).

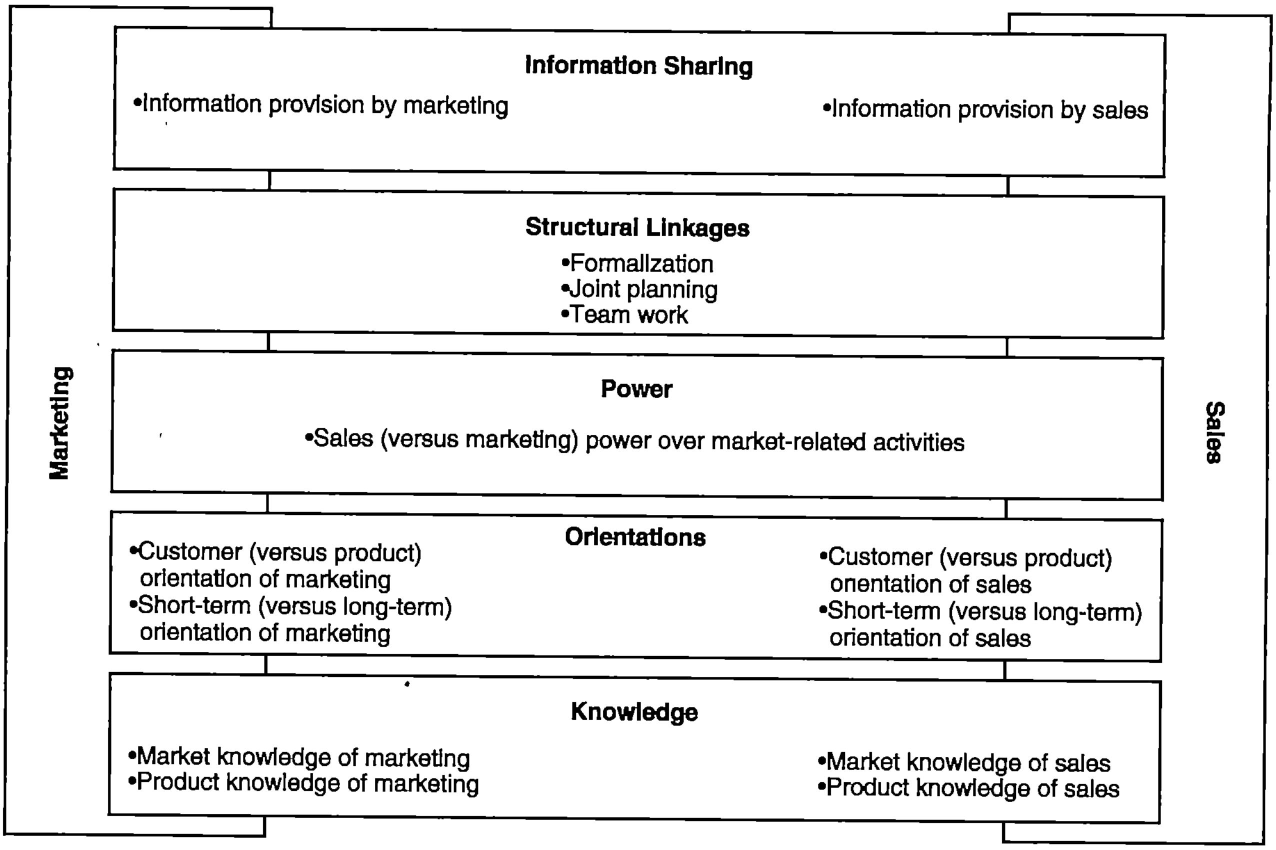
The third domain is power. It reflects how the influence over market-related activities is divided among organizational subunits (Homburg, Workman, and Krohmer 1999). This domain is particularly important to the taxonomy because extant research suggests a great deal of variation between firms. In some firms, sales units are dominant, and in others, marketing units are dominant (Piercy 1986; Workman, Homburg, and Gruner 1998).

The fourth conceptual domain groups various orientations. The orientations of an organizational unit pertain to the time horizon (short-term versus long-term) and objects (e.g., customers versus products) at which it arrays its activities (Lawrence and Lorsch 1969). For example, the literature has characterized sales as short-term oriented (Lorge 1999). In general, literature on interfunctional interfaces of

---

[1] Although the conceptual dimensions in Figure 1 do not represent a linear hierarchy, they are not completely unrelated. For example, from an organizational design perspective (Galbraith 1973; Lawrence and Lorsch 1969), differences between the orientations and the knowledge of M & S are viewed as differentiation, whereas structural linkages and information sharing are viewed as integration mechanisms. To close orientation and knowledge gaps between M & S, the literature recommends the use of integration mechanisms (Rouziès et al. 2005). Thus, conceptually, there is an interdependent relationship between the domains: More differentiation gaps require a higher use of integration mechanisms, whereas a higher use of integration mechanisms contributes to eliminating the differentiation gaps. The interdependent nature of the domains is consistent with our configurative analysis of construct patterns and speaks against specifying one-directional hypotheses for bivariate relationships.

FIGURE 1
Conceptual Model of M & S Interfaces


Marketing
Sales
Information Sharing
•Information provision by marketing
•Information provision by sales
Structural Linkages
•Formalization
•Joint planning
•Team work
Power
•Sales (versus marketing) power over market-related activities
Orientations
•Customer (versus product) orientation of marketing
•Short-term (versus long-term) orientation of marketing
•Customer (versus product) orientation of sales
•Short-term (versus long-term) orientation of sales
Knowledge
•Market knowledge of marketing
•Product knowledge of marketing
•Market knowledge of sales
•Product knowledge of sales

marketing has identified differing orientations as a major stressor to interfunctional cooperation (Griffin and Hauser 1996). The importance of functional orientations has also been noted in pioneer work on the M & S interface (Cespedes 1995).

The fifth domain is knowledge. This refers to the level of expertise in an organizational unit. The importance of knowledge to interdepartmental interfaces has been pointed out by research on departmental thought worlds. Knowledge differences can hamper communication and cause an interpretive barrier (Dougherty 1992; Frankwick et al. 1994; Workman 1993). Existing knowledge functions as a lens and affects how new information is weighted in making decisions (Brunswik 1955; Mitroff 1974).

Identifying Specific Constructs Within the Domains

Information sharing. The literature has conceptualized numerous constructs in the context of information sharing—for example, information transmission, intelligence dissemination, and information utilization (Jaworski and Kohli 1993; Moorman 1995). Of these constructs, information transmission, or intelligence dissemination, is conceptually the closest to interdepartmental information sharing. We view information transmission, or dissemination, as the key construct because it is a necessary condition to the use of information. Because information transmission, or generation, is a bilateral issue, we conceptualize information sharing with two variables: We define "information provision by marketing" as the extent to which marketing people supply salespeople with needed data and "information provision by sales" as the extent to which salespeople supply marketing people with needed data.

Structural linkages. Previously, we defined "structural linkages" as the extent to which formal horizontal platforms or channels are established for interdepartmental activities. The classical horizontal cooperation platform is "teamwork," which we define as the extent to which market-related activities are jointly developed and executed by members of M & S (Cespedes 1996). Another structural linkage that has received particular attention in the context of M & S is joint planning (Piercy 1989; Strahle, Spiro, and Acito 1996). We define "joint planning" as the extent to which M & S codevelop objectives, budgets, and activities. Finally, in the context of structural linkages, formalization is the most frequently studied dimension of the marketing organization (Dastmalchian and Boag 1990; Ruekert, Walker, and Roering 1985). We define "formalization" as the extent to which cooperation between M & S is structured by guidelines.

Power. The literature has used different approaches to conceptualize and operationalize power in an organizational context (Gaski 1987). One approach is based on the outcomes of power (Dahl 1957; Emerson 1962). A second approach is based on the sources of power (French and Raven 1959; Hunt and Nevin 1974; Lusch and Brown 1982). A third approach is based on decision areas over which power is exerted (Enz 1989; Gaski and Nevin 1985; Hinings et al. 1974). To assess interdepartmental power relations, marketing organization literature has typically used the third approach, measuring power as an index that assesses whether the relative power over market-related activities is concentrated in one or more departments (Homburg, Workman, and Krohmer 1999; Piercy 1989). We define "sales (versus marketing) power" as an index that aggregates the extent to which specific market-related activities are influenced by sales rather than by marketing. Through this power measure, we also have a proxy for the variation in the reporting lines of the marketing organization (Homburg, Workman, and Jensen 2000; Workman, Homburg, and Gruner 1998). Design variations ultimately manifest in varying degrees of power over marketing activities.

Orientations. The conceptual idea of departmental orientations has a long tradition in organizational research. The classical work of Lawrence and Lorsch (1969) winnows several facets of departmental orientations. The two facets that have been picked up the most frequently by later interface research are goal orientation and time orientation (Dougherty 1992; Griffin and Hauser 1996). These two facets are also reflected by managerial reports of the M & S interface (Cespedes 1995; Lorge 1999). It has been claimed that marketing focuses on products, whereas sales focuses on accounts. Our taxonomy shows whether this is the case for all companies or whether there also exist configurations of M & S with a customer-focused marketing and a product-minded sales force. We define "customer (versus product) orientation of marketing/sales" as the extent to which the activities of marketing/sales are guided by customer-related rather than product-related strategies, plans, and performance evaluations.

In addition, managerial reports claim that sales is rather short-term oriented and marketing is rather long-term oriented (Cespedes 1995; Lorge 1999). We inquire whether this applies across the board or whether some configurations have different time orientations. We define the "short-term (versus long-term) orientation of marketing/sales" as the extent to which activities of marketing/sales are guided by immediate action rather than by extensive planning.

Knowledge. Literature on departmental thought worlds has identified key constructs within the domain of knowledge. Dougherty (1992, p. 179) speaks of "technology-market knowledge." Whereas market knowledge pertains to the external environment of the organization, technology pertains primarily to the internal environment. Building on this dichotomy, we define "market knowledge of marketing/sales" as the extent to which a typical employee in marketing/sales is knowledgeable about customers and competitors, and we define "product knowledge of marketing/sales" as the extent to which a typical employee in marketing/sales is knowledgeable about products and internal processes.

Outcome Variables and Control Variables as Additional Descriptors

In addition to the cluster variables, we explore three outcome variables and three control variables. These variables are purely descriptive and do not enter the cluster analysis. Our first outcome variable characterizes the state of M & S cooperation. The literature often uses the integration construct (Kahn 1996; Leenders and Wierenga 2002). It was originally defined as "the quality of the state of collaboration that exists among departments that are required to achieve unity of effort by the demands of the environment" (Lawrence and Lorsch 1969, p. 11). However, Lawrence and Lorsch (1969, p. 11) note the ambiguity of the term "integration": "While we will be using the term 'integration' primarily to refer to this state of interdepartmental relations, we will also, for convenience, use it to describe both the process by which this state is achieved and the organizational devices used to achieve it." We define the "quality of cooperation between M & S" as the extent to which there is a state of collaboration between M & S that is characterized by unity of effort and harmony. We prefer the term "quality of cooperation" to the term "integration" because we refer to the state of interdepartmental relations, not to the integration process and the organizational devices.[2]

Our second and third outcome variables extend beyond the relationship of M & S and characterize the performance of the business unit (or company, if the organization has only one business unit). Typically, M & S are embedded in a business unit along with functions such as R&D, manufacturing, finance, and so forth. It is important to understand whether the M & S interface affects the performance of the whole. We emphasize that much prior literature on marketing's interfaces has merely considered the quality of cooperation between M & S and has not taken the performance of the whole business unit into account (De Ruyter and Wetzels 2000; Kahn, Reizenstein, and Rentz 2004; Maltz and Kohli 1996, 2000). We define "market performance" as the extent to which the business unit/company in which M & S are embedded achieves better market-related outcomes than its competitors with respect to metrics such as customer satisfaction and loyalty, new customer acquisition, market share, and so forth. We define "profitability" as the return on sales (ROS) (operating profit divided by operating revenues).

When we explore whether the configurations of M & S differ in regard to organizational outcomes, we control for three context variables. First, internal change has been demonstrated to encumber internal working relationships and attitudes (Reilly, Brett, and Stroh 1993), as well as marketing's interfunctional relationships (Maltz and Kohli 2000; Maltz, Souder, and Kumar 2001). Thus, we define "internal dynamism" as the extent to which organizational structures, processes, leadership, and strategy within the organization change frequently. Second, market turbulence and technological turbulence have been shown to influence the market performance of the firm (Kirca, Jayachandran, and Bearden 2005). Thus, we define "environmental dynamism" as the extent to which competitive activities, customer needs, and technology in the market change frequently. Third, we control for the influence of industry because industries differ by their ROS level.

---

[2] Our construct differs from the integration concept proposed by Song, Xie, and Dyer (2000), who specify three facets of integration: level of cross-functional involvement, quality of cross-functional information, and harmony of cross-functional relationships. The content of their measure of level of cross-functional involvement is reflected in our power domain, and their measure of quality of cross-functional information is reflected in our information-sharing domain. We view these more as determinants of cooperation quality. Our measure of quality of cooperation is closest to their measure of harmony of cross-functional relationships, but our measure also emphasizes the original "unity of effort" that Lawrence and Lorsch (1969) propose.

Methodology

General Approach to Configurative Analysis

Configurations have a long history in the study of organizations (Carper and Snizek 1980; Ketchen and Shook 1996). Two main approaches to configurations are used in the literature: The first approach analyzes organizational performance as a function of organizational fit with a contingency, typically structural fit with a strategic contingency (Doty, Glick, and Huber 1993; Venkatraman 1989). Fit is modeled as the proximity to an "ideal type" for each contingency. The most frequently studied contingencies are Miles and Snow's (1978) strategic types. For example, Vorhies and Morgan (2003) assess marketing organization fit with business strategy. This "fit approach" to configurations is confirmatory and requires a substantial amount of prior knowledge and theory about the subject matter. In contrast, the second approach to configurations is exploratory and demands little prior knowledge on the subject (Ketchen, Thomas, and Snow 1993). It uses cluster analyses to group organizations empirically. Our research follows the "classification approach" because the M & S interface is largely unexplored.

Data Collection Procedure

We conducted a large-scale mail survey of seven industry sectors in Germany (Table 3). We selected these sectors on the basis of their macroeconomic importance. By including consumer goods, industrial goods, and services, we ensured a variety of market settings. For example, prior research has noted differences between the M & S interface in consumer goods versus industrial firms (Cespedes 1993; Workman 1993). We obtained a random sample of SBUs from a commercial list provider. All SBUs came from different firms. Telephone calls verified that the SBU had both a marketing and a sales subunit and identified informants with an overview of both subunits. Our sample included executives from 1700 SBUs.

All respondents were addressed by a personalized letter and follow-up telephone calls. The respondents knew that their answers would be treated confidentially. Respondents were asked to refer to their SBU or, if their company only had one SBU, to their company. All respondents were asked to answer both the marketing-related and the sales-related questions. Given our prior identification of SBUs with separate M & S units, respondents were familiar with both terms and were asked to refer to the units labeled M & S. To ensure the appropriateness of the respondents, two items at the end of the questionnaire asked how competent the respondents were to answer both the marketing-related and the sales-related questions and how strongly the respondent was involved in the cooperation between M & S. We discarded 38 returned questionnaires because one of the items was rated lower than four on a five-point scale (five indicating high competence). As a result, we obtained 337 completed and usable questionnaires. The average job experience of these people in M & S is 16.6 years, with a median of 14 years. Our response rate is 20%. Although nonresponse is always a concern in survey research, the response rate is within the range of comparable research (Harzing 1997; Olson, Slater, and Hult 2005). We controlled for nonresponse bias by comparing construct means for early and late respondents (Armstrong and Overton 1977). Of 18 perceptual variables, only 1 (market knowledge of marketing) showed significant ($p < .05$) differences between early and late respondents. This speaks against a systematic bias.

Because the cluster variables and the outcome measures stem from the same person, which may create a common method bias (Podsakoff et al. 2003), we collected financial performance data from independent sources to validate the market performance measure. We retrieved annual reports from two financial databases and the firms' Web sites. We used segment reports when available. We obtained the ROS (operating profits divided by operating revenues) of the year in which the survey was administered and for the three subsequent years. Overall, these procedures yielded financial data for 185 of the 337 data sets.[3] We compared construct

TABLE 3
Sample Composition

Industry		Position of Respondent		Number of Employees in Business Unit/Company	
Financial services	26%	Head of marketing	46%	<200	38%
Consumer packaged goods	19%	Head of sales	36%	200–399	14%
Utilities	12%	General manager/director	4%	400–999	17%
Chemical/pharmaceutical	12%	Head of marketing communication	4%	1000–2999	18%
Automotive	11%	Head of product management	4%	3000 or more	13%
Machinery	10%	Other	6%		
Electronics	5%				
Other	5%				

---

[3] We could not obtain financial performance data for many family-, foundation-, and state-owned companies, because public-

means for data sets with and without financial performance. None of the 18 perceptual variables showed significant ($p <$ .05) differences. In particular, the means of market performance (3.53 versus 3.46, t = .84) and cooperation quality (3.62 versus 3.53, t = .96) were not significantly different. Thus, this test speaks against the existence of a major availability bias.

Measurement Procedure

We use two types of perceptual measures: reflective multi-item scales and formative multi-item indexes. If observed variables are interchangeable manifestations of an underlying construct, a reflective measurement model is appropriate (Bagozzi and Baumgartner 1994). In this study, we apply a reflective model to measure constructs such as orientations, formalization, teamwork, internal and environmental dynamism, and quality of cooperation. On the contrary, if a construct is a summary index of observed variables that define and determine the construct without necessarily being correlated, a formative measurement model is appropriate (Jarvis, MacKenzie, and Podsakoff 2003). We contend that a formative model is appropriate to measure knowledge and power; being knowledgeable in one area does not entail being knowledgeable in another, especially because specialization is a hallmark of organizations. In regard to power, if sales dominates price-related tasks, that does not inevitably mean that sales dominates market research tasks as well (Homburg, Workman, and Krohmer 1999; Kotler, Rackham, and Krishnaswamy 2006). Treating power as an index has a long tradition in the organizational literature (Eisenhardt and Bourgeois 1988; Hinings et al. 1974).

We designed the questionnaire after an extensive literature review and qualitative interviews with at least one manager from each of the industries. A draft was pretested and refined with 20 additional managers. The Appendix provides scale items, properties, and sources. We assessed reliability and validity of the reflective measures with multifactorial confirmatory factor analysis. Each construct manifests a composite reliability of at least .6 (Bagozzi and Yi 1988). Average variance extracted is higher than the squared phi coefficient for any pair of two latent variables, which supports the discriminant validity of the reflective measures (Fornell and Larcker 1981).

In constructing formative indexes, four critical issues are content specification, indicator specification, indicator collinearity, and external validity (Diamantopoulos and Winklhofer 2001). In the "Conceptual Model" section, we specified the content of power as the aggregate influence over market-related activities. On the basis of the literature review and the interviews, we specified seven activity indicators (see the Appendix). We specified the content of market knowledge as knowledge about customers and competitors. The indicators comprise the external actors in Ohmae's strategic triangle. We specified the content and the indicators of product knowledge as knowledge about the product outcome and internal processes of creating and delivering it. After scrutinizing indicator collinearity for the three indexes, we did not find collinearity to be a problem. The maximum squared bivariate interindicator correlation is .16 (Cohen et al. 2003). To assess the external validity of the power indicators, we followed a procedure similar to that of Coviello, Winklhofer, and Hamilton (2006); that is, we correlated each of the seven indicators to an external constant-sum measure of overall sales versus marketing power. Four correlations are significant at a .01 level, and three correlations are significant at a .05 level, which supports the external validity of the power measure. We aggregated the individual indicators to the formative indexes using equal weights (Dillon 2001).

As Table 3 shows, the sample cumulates respondents from marketing and respondents from sales. This makes it necessary to probe whether the structural patterns in the data set differ between marketing respondents and sales respondents: We tested whether the correlation matrix of the indicator variables differs between M & S respondents.[4] The null hypothesis that variable correlations of the sales and the marketing respondents are equal cannot be rejected at a 5% significance level. This test represents strong evidence that responses from M & S do not differ and that pooling the two groups is justified. In the "Results" section, we report an additional test, in which we assess whether some of the emerging clusters reflect primarily responses from sales and other clusters reflect primarily responses from marketing. Again, we find no evidence that the functional membership of the respondent biases the empirical results.

---

ity requirements are less comprehensive in Europe than in the United States. An additional challenge arises in the case of very large companies. For such enterprises, comprehensive financial data are available on the level of the entire corporation, but limited data are available on the level of divisions, or segments. However, the divisions are still billion-dollar businesses and can be composed of dozens of business units. The business units are medium-sized enterprises, but their financial data are not publicly available. To verify whether the database ROS measures the performance of the responding business unit and not the performance of a much larger organizational division, we compared self-stated revenues from the survey with the revenues in the financial report. In some cases, the business unit that responded to our survey represented only a small fraction of the revenues in the financial report because the financial report combines the revenues from several business unit (whose individual revenues are not public). In these cases, when database revenues and profits are dominated by other business units of the firm rather than the business unit responding to the survey, the database ROS could not be used as a performance measure of the responding business unit.

[4] Using the two largest groups in Table 3, the "Head of Sales" (n = 120) and the "Head of Marketing" (n = 155), we estimated a two-group LISREL measurement model with the indicator variables of all constructs. We modeled each of the 52 indicator variables as a single-item exogenous variable measured without error. We conducted a chi-square difference test between an unconstrained model that freely estimated the phi matrices in the marketing and the sales group (i.e., the estimated correlations of indicators) and a model that constrained the phi matrices in both groups to be equal. The equality model fits the data well (root mean square error of approximation = .02, comparative fit index = .95, and nonnormed fit index = .91). The chi-square difference between the two models is 1454.9 (1378 = d.f.); thus, $p > .05$. We cannot reject the null hypothesis of equal phi matrices.

Taxonomic Procedure

We took a three-stage clustering approach, building on procedures that Bunn (1993), Cannon and Perreault (1999), and Homburg, Workman, and Jensen (2002) use. The three core issues in clustering are determining the number of clusters, assigning observations to clusters, and assessing the stability of cluster assignments. We conducted all analyses with SAS 9.1.

To determine the appropriate number of clusters, we used the cubic clustering criterion (Sarle 1983) and the pseudo-t^2 index (Duda and Hart 1973) in combination with the hierarchical clustering algorithm that Ward (1963) developed. In Milligan and Cooper's (1985) comparative study of 30 methods for estimating the number of population clusters, cubic clustering criterion and pseudo-t^2 were among the top-performing criteria. Because Ward's algorithm is sensitive to scaling and outliers, we standardized the clustering variables on the basis of their range (Milligan and Hirtle 2003). In addition, we identified 15% of the observations as outliers and removed them from the data set.[5] This is consistent with Punj and Stewart (1983, p. 143), who note that outliers can hamper algorithm performance when more than 90% of all observations are fed into the cluster algorithm. Finally, clustering 20 randomly selected subsamples from the remaining data, each containing two-thirds of the sample, we found strong support for a five-cluster solution.

To assign observations to clusters, we clustered the complete sample by a hybrid approach (Punj and Stewart 1983) that combined Ward's method with the k-means method. Simulation studies on the performance of clustering algorithms demonstrate that partitioning methods (e.g., k-means) yield excellent results if given a reasonable starting solution (Milligan and Cooper 1987). Using Ward's method to compute a starting solution for k-means has been shown to be a powerful combination (Helsen and Green 1991) and has been recommended by Punj and Stewart (1983) and Arabie and Hubert (1994).

To assess the stability of the cluster assignment, we used the cross-validation procedure that McIntyre and Blashfield (1980) propose. Milligan and Cooper (1987, p. 351) deem this procedure to be "an excellent strategy for establishing the generalizability of a cluster analysis." We randomly split the sample in halves and applied the hybrid clustering procedure to each half. Then, we assigned each object in the second half to the nearest cluster centroid obtained from the first half (based on the squared Euclidean distance). As a result, we obtained two cluster assignments for each object in the second half and compared their consistency using Rand's (1971) index, which counts not only positive but also negative coincidence. It is scaled between 0 and 1 (1 indicates perfect stability). The cross-validation yields a Rand index of .8.

Results

The last step toward the taxonomy is to validate whether the clusters have meaningful interpretations (Rich 1992). Table 4 shows the cluster means for each of the 14 cluster variables. It also shows descriptive variables that explore the outcomes of the clusters, their representation in various industries, company size categories, respondent categories, and contexts. It is worth reiterating that the clusters were formed only on the basis of the 14 constructs from the five interface domains. The descriptive variables were not used as active cluster variables. Thus, we explore whether the clusters differ in regard to the descriptive variables. Specifically, we tested whether cluster membership is associated with the respondent's functional background, company size, or industry. We find no significant indication that some clusters reflect respondents primarily from sales and others primarily from marketing ($\chi^2 = 5.13$, d.f. = 4, $p = .27$), which adds confidence to the key informant data. We also find no significant association between cluster membership and company size ($\chi^2 = 16.02$, d.f. = 16, $p = .45$). However, we find a significant association between cluster membership and industry ($\chi^2 = 84.63$, d.f. = 3 2, $p < .01$), which we expand on in the description of the individual clusters. Following the interpretation steps that Bunn (1993) suggests, we compared the cluster means on the continuous variables, using Waller and Duncan's (1969) k-ratio t-test. Cluster means carrying the same superscript do not differ at a 5% significance level. Finally, we translated the statistical ranges into verbal descriptions (see Table 5).

Interpretation of Clusters

In interpreting the clusters, we assign names to the approaches. Although such names tend to oversimplify, they point out empirically distinct aspects of different approaches and facilitate the discussion of the results. For illustrative purposes, we also indicate prototypical firms that are the closest to the centroid of each cluster.

Cluster 1 ("Ivory Tower"). The sales unit is in the lowest range of market knowledge and is the most product focused of all clusters. It also has a rather short-term focus. Thus, sales appears to have an operative rather than a strategic mind-set. It seems that the sales unit has a product-driven, operative selling role. This contrasts sharply with the marketing unit, which has a rather high customer focus and a medium time orientation. Whereas the high customer orientation could put marketing in a position to counterpoise the operative sales force, marketing has one problem: It is out of touch with the market and the products. Marketing's level of product and market knowledge is the lowest of all clusters. In addition, there is not much information sharing between sales and marketing and little joint planning. Marketing's role appears to be an isolated caller for more customer orientation. Overall, the most obvious characteristic of this cluster is that both M & S seem to lack an under-

---

[5] We identified outliers using a dual criterion We used nearest-neighbor (single-linkage) clustering to mark observations that were joined the last in the agglomeration tree. We also used PROC FASTCLUS to compute a 20-cluster solution and marked observations that formed their own, small cluster detached from the other observations. If both criteria were in agreement, the observation was identified as an outlier.

TABLE 4
Statistical Cluster Description

		Total	1	2	3	4	5
Percentage of Observations		**100% (n = 278)**	**28% (n = 77)**	**18% (n = 51)**	**14% (n = 38)**	**17% (n = 48)**	**23% (n = 64)**
Functional background of informants	Head of marketing	100%	29%	16%	12%	16%	27%
	Head of sales	100%	34%	22%	13%	15%	16%
Number of employees in business unit/company	<200	100%	30%	13%	16%	12%	29%
	200–399	100%	31%	16%	5%	24%	24%
	400–999	100%	25%	23%	10%	23%	19%
	1000–2999	100%	22%	24%	15%	17%	22%
	3000 or more	100%	28%	22%	19%	19%	12%
Industry	Financial services	100%	44%	9%	9%	9%	29%
	Consumer packaged goods	100%	26%	39%	9%	14%	12%
	Machinery	100%	32%	8%	25%	21%	14%
	Automotive	100%	23%	23%	25%	10%	19%
	Electronics	100%	25%	8%	8%	42%	17%
	Chemical/pharmaceutical	100%	15%	12%	6%	41%	26%
	Utilities	100%	19%	13%	25%	6%	37%
Information sharing	Information provision by marketing		3.50^{c}	4.34^{a}	3.99^{b}	3.60^{c}	4.06^{b}
	Information provision by sales		$3.18^{b,c}$	4.22^{a}	3.25^{b}	2.96^{c}	3.41^{b}
Structural linkages	Formalization		2.73^{c}	3.52^{a}	1.97^{d}	3.03^{b}	3.00^{b}
	Joint planning		3.75^{c}	4.50^{a}	2.88^{d}	3.78^{c}	4.23^{b}
	Teamwork		3.54^{b}	4.17^{a}	2.64^{c}	3.55^{b}	4.02^{a}
Power	Sales (versus marketing) power		$3.30^{b,c}$	3.15^{c}	3.73^{a}	2.67^{d}	3.41^{b}
Orientations	Marketing's customer (versus product) orientation		2.95^{b}	2.59^{c}	2.90^{b}	1.91^{d}	3.86^{a}
	Marketing's short-term (versus long-term) orientation		2.77^{b}	2.28^{c}	3.30^{a}	2.46^{c}	2.98^{b}
	Sales' customer (versus product) orientation		2.78^{c}	3.26^{b}	3.11^{b}	3.31^{b}	4.01^{a}
	Sales' short-term (versus long-term) orientation		3.48^{b}	2.73^{d}	$3.39^{b,c}$	3.91^{a}	3.15^{c}
Knowledge	Marketing's market knowledge		3.09^{b}	3.64^{a}	3.76^{a}	3.27^{b}	3.82^{a}
	Marketing's product knowledge		3.44^{d}	4.57^{a}	3.95^{c}	4.28^{b}	3.77^{c}
	Sales' market knowledge		3.83^{c}	4.54^{a}	4.46^{a}	3.76^{c}	4.18^{b}
	Sales' product knowledge		3.51^{c}	4.08^{a}	4.13^{a}	3.04^{d}	3.88^{b}
Outcomes (descriptive)	Cooperation quality		3.36^{c}	4.23^{a}	3.33^{c}	3.25^{c}	3.77^{b}
	Market performance		$3.37^{b,c}$	3.87^{a}	3.21^{c}	3.50^{b}	3.51^{b}
	Profitability$_{t+0}$		$.03^{b}$	$.06^{a}$	$.03^{b}$	$.03^{b}$	$.04^{b}$
	Profitability$_{t+1}$		$.03^{b}$	$.06^{a}$	$.04^{b}$	$.03^{b}$	$.03^{b}$
	Profitability$_{t+2}$		$.02^{b}$	$.07^{a}$	$.04^{b}$	$.01^{b}$	$.04^{b}$
	Profitability$_{t+3}$		$.03^{b}$	$.09^{a}$	$.04^{b}$	$.00^{b}$	$.04^{b}$
Context variables (descriptive)	Internal dynamism		2.69^{a}	2.62^{a}	2.73^{a}	2.83^{a}	2.72^{a}
	Environmental dynamism		2.90^{a}	3.07^{a}	2.50^{b}	$2.75^{a,b}$	$2.80^{a,b}$

Notes: Reported values are mean values if not indicated otherwise. In each row, cluster means that have the same superscript are not significantly different ($p < .05$) on the basis of Waller and Duncan's (1969) multiple-range test. Means in the highest bracket are assigned the superscript "a," means in the next lower bracket are assigned the superscript "b," and so forth.

standing for products and market needs. Lorge (1999) metaphorically refers to such a situation as being locked in the Ivory Tower.

This configuration of M & S is common in financial services and the machinery industry. Three typical cluster members are a travel insurance company, a producer of plumbing and radiant heating solutions, and a producer of professional catering technology. In Germany, banks have gone through a severe crisis. They are undertaking massive efforts to improve the customer orientation of their sales force and to train their frontline employees in more advanced capital market instruments. During the expert

TABLE 5
Verbal Cluster Description

		Ivory Tower	Brand-Focused Professionals	Sales Rules	Marketing-Driven Devil's Advocacy	Sales-Driven Symbiosis
Interpretation	Marketing role	Isolated caller for customer orientation	Expert in leading role	Operative support	Long-term product voice	Market expert
	Sales role	Product-driven, operative selling	Congenial counterpart of marketing	Dominant product expert	Operative customer voice	Dominant product expert
Industries in which cluster is strongly represented		Financial Services, Machinery	Consumer Goods	Machinery, Automotive	Chemical, Electronics	Utilities
Information sharing	Information provision by marketing	Low	High	Medium	Low	Medium
	Information provision by sales	Low–medium	High	Medium	Low	Medium
Structural linkages	Formalization	Moderately low	High	Low	Moderately high	Moderately high
	Joint planning	Moderately low	High	Low	Moderately low	Moderately high
	Teamwork	Medium	High	Low	Medium	High
Power	Sales (versus Marketing) Power	Moderately low–moderately high	Moderately low	High	Low	Moderately high
Orientations	Marketing's customer (versus product) orientation	Moderately high	Moderately low	Moderately high	Low	High
	Marketing's short-term (versus long-term) orientation	Medium	Low	High	Low	Medium
	Sales' customer (versus product) orientation	Low	Medium	Medium	Medium	High
	Sales' short-term (versus long-term) orientation	Moderately high	Low	Moderately low–moderately high	High	Moderately low
Knowledge	Marketing's market knowledge	Low	High	High	Low	High

TABLE 5
Continued

		Ivory Tower	Brand-Focused Professionals	Sales Rules	Marketing-Driven Devil's Advocacy	Sales-Driven Symbiosis
	Marketing's product knowledge	Low	High	Moderately low	Moderately high	Moderately low
	Sales' market knowledge	Low	High	High	Low	Medium
	Sales' product knowledge	Moderately low	High	High	Low	Moderately high
Comparison with conceptual classifications	Day (1999)		Marketing as central guidance function	Marketing as subordinate function	Marketing as functional fiefdom	
	Kotler, Rackham, and Krishnaswamy (2006)			Marketing as an adjunct to sales	Marketing as an independent player	
	Webster (1997)		Marketing integrated with other business functions		Marketing as a distinct function	

interviews, we observed a potential reason marketing is detached from the market in these industries. Companies sometimes hire marketing professionals from consumer goods companies to inspire marketing thought in the organization. These people are highly customer oriented but cannot really make themselves heard, because they know too little about the specificities of the market and the products.

Cluster 2 ("Brand-Focused Professionals"). This cluster manifests the highest levels of formalization, joint planning, teamwork, and information sharing. In addition, both M & S have the highest levels of market and product knowledge and the most long-term orientation of all clusters. Thus, the cooperation between M & S appears to be structured and professional in this cluster. Notably, marketing is rather powerful and product driven in this cluster. This might be because this cluster represents the typical symbiosis of M & S in the consumer packaged goods industry. This industry is characterized by a strategic focus on brands and by powerful brand managers, but in recent years, sales and key account managers have gained status in consumer goods companies as customer business development, efficient consumer response, and category management are implemented (Cespedes 1993; Homburg, Workman, and Jensen 2000). Our study shows that marketing is still in the lead role and that sales has become a congenial counterpart. Two prototypical cluster members are a producer of soft drinks and a producer of chocolate.

Cluster 3 ("Sales Rules"). Sales dominates the stage in this cluster. Sales is the most powerful of all clusters and among the most knowledgeable, with respect to both products and markets. Notably, marketing has rather low product knowledge, so the role of the product expert apparently accrues to sales in this cluster. Marketing has the highest short-term orientation of all clusters. This can be interpreted such that marketing plays the role of operative support to the dominant sales unit. We do not find much structured cooperation in this cluster—formalization, joint planning, and teamwork are the lowest of all clusters—which reinforces the impression that marketing is little more than an appendix to the sales force. Two typical cluster members are a producer of wind turbines and a local electric utility. Wind turbines are a project-based business with a small number of (often public) customers. Some local electric utilities have only recently been deregulated after having been local monopolists with little need for marketing. This could explain why the role of marketing is so small in these companies. In general, this cluster is common in the machinery and automotive industry.

Cluster 4 ("Marketing-Driven Devil's Advocacy"). This is a cluster of opposite orientations. Sales has the most short-term, operative focus of all clusters, whereas marketing has a long-term, strategic focus. Marketing has the strongest product orientation of all clusters and rather high product knowledge, whereas sales has the lowest product knowledge of all clusters. At the same time, cooperation between M & S is rather formalized. It seems that this configuration is an institutionalized controversy: In a fairly formalized process, the contrasting perspectives of marketing as a strategic product voice and sales as an operative customer voice are confronted for optimally balanced decisions. Literature on strategy making refers to such decision processes as dialectical inquiry or devil's advocacy (Dooley and Fryxell 1999; Mason and Mitroff 1981). The final word in these discussions belongs to marketing, which is the most powerful of all clusters. The dominance of product-focused marketing fits with the observation that this is the typical configuration in the chemical and electronic industry. These industries are characterized by capital-intensive production assets. Capacity utilization drives many market-related decisions and is closely observed by product managers. This could explain why marketing is long-term and product oriented but has low market knowledge in this cluster. Marketing seems to be technical and inward looking. In other words, being product oriented is strongly associated with being production oriented in this cluster. Three typical cluster members are a business unit of a producer of fertilizers, a major airfreight company, and a European plant of a U.S. car manufacturer.

Cluster 5 ("Sales-Driven Symbiosis"). This is a cluster of complementary skills. Marketing's product knowledge is rather low, whereas sales' product knowledge is rather high. Marketing's market knowledge is high, whereas sales' market knowledge is medium. Both M & S are highly customer focused, and the level of teamwork is in the highest range of all clusters. In addition, formalization and joint planning are rather high, which suggests a structured cooperation. Sales is the more powerful player in this cooperation. We chose the term symbiosis for this cluster to emphasize the interdependent teamwork of two departments with complementary skills. This configuration is typical for utilities. The two most typical cluster members are a local electrical utility firm and a business unit of one of Europe's largest utilities. Because utilities are characterized by long-term contracts, it is plausible that sales is long-term oriented in this cluster.

Exploration of Outcomes

The bottom of Table 4 shows the outcomes achieved by the five clusters. Cooperation quality and market performance are retrospective perceptual measures that reflect the situation up to the survey. In addition, we collected objective profitability data for the years following the survey to address causality issues. We first comment on the outcomes achieved by each cluster (see Table 4) and then make more general comments on the findings. The Ivory Tower cluster is in the lowest range of clusters for all three outcome variables. The profitability of this cluster is consistently in the lowest category. This is not surprising given the low levels of market knowledge, information sharing, and structural linkages. In contrast, the Brand-Focused Professionals cluster achieves the highest levels of cooperation quality, market performance, and objective financial performance. We emphasize that this cluster's profitability is consistently in the highest bracket for several years following the survey. The high levels of information sharing, knowledge, and structural linkages in this cluster pay off. Sales Rules is one of the weaker configurations. All three outcomes are in the lowest range of all clusters. It seems that this cluster's most noticeable features, the low extent of structural linkages and the highly unilateral power concentration, are not a promising solution for the M & S interface. Marketing-Driven Devil's Advocacy has the lowest cooperation quality of all clusters. This is consistent with the observation that it is the cluster of clashing orientations, with sales as the operative customer voice and marketing as the strategic product voice. It seems that this devil's advocacy, by contrasting alternative routes, can produce fairly good decisions. The market performance of this cluster is in the middle range. This is consistent with research on the benefits of strategic dissensus (Bourgeois 1985). However, in following years, the profitability of this cluster is consistently in the lowest range. Finally, Sales-Driven Symbiosis is in the second-highest performance bracket in regard to cooperation quality, market performance, and profitability.

What the two most successful configurations (Brand-Focused Professionals and Sales-Driven Symbiosis) have in common is an intense use of structural linkages, high market knowledge within the marketing unit, and a long-term orientation of the sales unit. These two clusters are characterized by a clear, but not extreme, power distribution between M & S. The pattern that pervades the less successful configurations is low levels of information sharing, structural linkages, and knowledge as well as extreme power distribution.

It is necessary to verify whether the between-cluster differences of the outcome variables remain significant when we control for context factors. We analyzed seven analysis of covariance (ANCOVA) models—one with cooperation quality as the dependent variable, one with market performance, and five with profitability at different time lags (Table 6). The ANCOVA models comprised four predictors: two class variables (cluster membership and industry) and two metric covariates (internal dynamism and environmental dynamism). As Table 6 shows, the performance impact of the M & S configuration remains highly significant for all outcomes even after the inclusion of control variables. We find a significant association of cluster membership with profitability not only in the year of the survey but also in subsequent years. We find the highest partial F value for cluster membership for a lag of +1 year after the survey. We emphasize that the model explains 35% of the variance of profitability in that year.[6]

To further corroborate trust in the effect of M & S configurations on profitability, we conducted a Granger-causality test (Dufour and Renault 1998; Granger 1969). The M & S configuration is assumed to Granger-cause profitability if the M & S configuration at $t + 0$ provides statistically relevant information about profitability in $t + n$ in the presence of profitability at $t + 0$. As the effect of configuration is significant for various time lags n, we find strong

---

[6] Variance explanation is even higher in years $t + 2$ and $t + 3$, but this is more of an industry effect than an effect of the M & S configuration.

TABLE 6
Exploratory ANCOVA Results for Cluster Performance

Dependent Variable		d.f.	Type III Sum of Squares	Mean Sum of Squares	F	R^2
Cooperation quality	Model	14	42.88	3.06	7.09***	.27
	Error	260	112.29	.43		
	M & S configuration	4	32.07	8.02	18.56***	
	Industry	8	5.11	.64	1.48	
	Market dynamism	1	2.32	2.32	5.37**	
	Internal dynamism	1	1.17	1.17	2.71	
Market performance	Model	14	20.88	1.49	3.97***	.18
	Error	259	97.40	.38		
	M & S configuration	4	5.94	1.49	3.95***	
	Industry	8	6.60	.83	2.20**	
	Market dynamism	1	2.02	2.02	5.37**	
	Internal dynamism	1	1.45	1.45	3.84*	
Profitability$_{t+0}$	Model	14	.0263	.0019	2.48***	.20
	Error	136	.1033	.0008		
	M & S configuration	4	.0098	.0024	3.22**	
	Industry	8	.0123	.0015	2.03**	
	Market dynamism	1	.0013	.0013	1.75	
	Internal dynamism	1	.0002	.0002	.29	
Profitability$_{t+1}$	Model	14	.0492	.0035	5.05***	.35
	Error	129	.0899	.0007		
	M & S configuration	4	.0113	.0028	4.07***	
	Industry	8	.0247	.0030	4.42***	
	Market dynamism	1	.0001	.0001	.18	
	Internal dynamism	1	.0004	.0004	.63	
Profitability$_{t+2}$	Model	14	.1369	.0098	5.60***	.39
	Error	123	.2147	.0017		
	M & S configuration	4	.0154	.0038	2.20*	
	Industry	8	.0916	.0114	6.56***	
	Market dynamism	1	.0003	.0003	.16	
	Internal dynamism	1	.0001	.0001	.05	
Profitability$_{t+3}$	Model	14	.3680	.0263	6.57***	.43
	Error	120	.4800	.0040		
	M & S configuration	4	.0456	.0114	2.84**	
	Industry	8	.2922	.0365	9.12***	
	Market dynamism	1	.0074	.0074	1.84	
	Internal dynamism	1	.0002	.0002	.05	

*$p < .10$.
**$p < .05$.
***$p < .01$.

support that the M & S configuration Granger-causes profitability.[7]

To conclude, on the one hand, the exploration of outcomes reveals one cluster (Cluster 2) whose profitability significantly differs from the others and two clusters (Clusters 2 and 5) whose cooperation quality and market performance differ from the others. This suggests that there are superior configurations for which companies should strive. On the other hand, we observe that several very different clusters achieve similar outcomes. This is consistent with the equifinality hypothesis of the configurative approach (Meyer, Tsui, and Hinings 1993).

---

[7] We first estimated a general linear model for profitability at t + 1 as the dependent variable with the M & S configuration at t + 0 and profitability at t + 0 as the two independent variables. A likelihood ratio test shows that the effect of the M & S configuration is significant ($\chi^2 = 13.73$, d.f. = 4, $p < .01$), even in the presence of the lagged profitability variable ($\chi^2 = 37.11$, $p < .01$). Using profitability at t + 2 as the dependent variable, we found both the effect of the M & S configuration ($\chi^2 = 14.07$, d.f. = 4, $p < .01$) and the effect of profitability at t + 0 ($\chi^2 = 4.95$, $p = .03$) to be significant again. For a lag of +3 years, the effects of both M & S configuration ($\chi^2 = 9.29$, d.f. = 4, $p = .05$) and profitability at t + 0 ($\chi^2 = 2.97$, $p = .08$) are weaker but still weakly significant.

Discussion

Comparison of Cluster Results with Conceptual Typologies in the Literature

We compare the empirical taxonomy with previous conceptual typologies in the literature. The bottom three rows of Table 5 juxtapose the conceptual types and the emergent clusters. From these, we draw three conclusions: First, there is a high consistency among the conceptual typologies. The literature essentially describes three types of M & S configurations. Second, all three conceptual types can be matched to empirical clusters in this study. Third, our study discovers two additional clusters that have not been discussed in the literature.

We now discuss each of the conceptual types. As a first type with a limited marketing role, both Day (1999) and Kotler, Rackham, and Krishnaswamy (2006) propose that, in some companies, marketing is a subordinate adjunct to sales. The finding of the Sales Rules cluster confirms this proposition. As a second and more pronounced marketing role, all three sources propose a distinct and independent marketing "fiefdom" that challenges the sales function and is characterized by palpable conflict between M & S. This proposition remarkably parallels the Marketing-Driven Devil's Advocacy cluster, which has the lowest cooperation quality of all clusters. As a third marketing role, Day suggests that "marketing as a central guidance function" is characterized by a strong marketing department, a highly market-oriented culture, and successful market leadership. In a similar vein, Webster (1997) suggests that "integrated marketing" is characterized by a strong marketing department, intense use of teams and other structural linkages, and a high level of success. The Brand-Focused Professionals cluster confirms these propositions. Overall, the finding that three of the five clusters are closely related to previous conceptual typologies supports the validity of the cluster results.

Two of the configurations have not been discussed in the literature before: the Ivory Tower and Sales-Driven Symbiosis clusters. This can be explained in several ways: First, we reiterate that the conceptual sources do not claim to present a complete typology of M & S configurations but rather to contrast the most salient types. Second, conceptual typologies tend to reflect equilibrium states of M & S that are stable for some time. In contrast, empirical taxonomies also capture intermediate stages of change processes that punctuate the equilibrium states (Tushman and Romanelli 1994). As we discussed previously, the isolated marketing function in the Ivory Tower cluster can be viewed as a transition step of a customer-oriented change process. Third, the three types that have been proposed by conceptual typologies are common in industries that rank at the top of business press attention: consumer goods, automotive, and electronics. In contrast, one of the additionally discovered clusters is typical of an industry that receives comparatively less coverage, namely, utilities. Fourth, the core focus of the extant literature is on marketing's role in the firm, not on sales. Not surprisingly, the conceptual types include the two clusters in which marketing's role is the strongest and the one cluster in which marketing's role is the weakest. In contrast, the two additional clusters represent shades of gray that are predominantly driven by sales. To conclude, the comparison shows the importance of validating and complementing conceptual typologies by broad-based empirical taxonomies.

Implications for Academic Research

This study contributes to academic marketing research in several ways. First, it is one of the first to consider the managerially important interface between M & S. This interface has long been neglected in the literature, but lately, it has been chosen as one of the topics of the American Marketing Association thought leadership forums (Rouziès et al. 2005). To the best of our knowledge, ours is the largest empirical database of the M & S interface so far, collecting questionnaires from several industries and more than 300 companies. This lends generalizability to the findings.

Second, this article makes a conceptual contribution. We identify dimensions of M & S configurations and develop a multidimensional model for studying interfaces. More specifically, we distinguish five domains: information sharing, structural linkages, power, orientations, and knowledge. Whereas information sharing, structural linkages, and power issues have received significant research attention, orientations and knowledge are comparatively less explored. The importance of studying orientations and knowledge has been emphasized in research on departmental thought worlds (Dougherty 1992; Frankwick et al. 1994). In addition, whereas prior research has studied selected domains, there has been little integrative research that considers all five domains simultaneously. Thus, a particular contribution of this article is its conceptual breadth. We believe that the multidimensional conceptualization is helpful not only to the study of the M & S interface but also to the study of other intraorganizational interfaces (e.g., sales and production, marketing and R&D) and even to the study of interorganizational interfaces (e.g., with marketing alliance partners or with advertising agencies).

Third, this article develops a taxonomy that describes which archetypical configurations of M & S occur in practice. The taxonomy reveals significant variation in the roles and cooperation of M & S. This finding challenges stereotypes that are often published about M & S. Caution is warranted when making general statements about marketing or sales. For example, in some configurations, marketing has more product knowledge than sales, whereas the converse applies to others. The taxonomy confirms three earlier conceptual typologies that have been proposed in the literature and identifies two additional configurations.

Fourth, this article has ramifications for the sampling frame of future studies on marketing organization. Some marketing organizations are marketing dominated, as the textbooks suggest, but other so-called marketing organizations are sales dominated. Researchers should make a conscious choice of the M & S configurations they study. If it is not possible to develop grand theories about marketing organization that apply to all configurations, it may be possible to develop midrange theories (Merton 1968; Pinder and Moore 1979) for some.

Fifth, this study explores the outcomes of M & S configurations. The exploratory results suggest that Brand-Focused Professionals and Sales-Driven Symbiosis clusters are more successful than other configurations. The more successful clusters are characterized by strong structural linkages and marketing's market knowledge. Note also that in the most successful cluster, the marketing function has a strong role. This exploratory finding supports prior research that has emphasized the importance of functional marketing (Moorman and Rust 1999). However, the finding that the Sales-Driven Symbiosis cluster is more successful than the Marketing-Driven Devil's Advocacy cluster in terms of cooperation quality and market performance also cautions not to exaggerate academic demands for making marketing the lead function.

Implications for Managerial Practice

The conceptual and empirical results have important implications for managerial practice. First, the taxonomy helps managers understand the specificities of M & S configurations in their own company and develop alternative solutions. We find that five configurations emerge from 14 cluster variables, which indicates that there are only a few "stable" and internally consistent configurations. However, the performance of these stable configurations differs. Brand-Focused Professionals and Sales-Driven Symbiosis clusters perform better than the rest with respect to outcomes, suggesting two target configurations for which managers should strive. In contrast, the feeble performance of the other clusters indicates traps that should be avoided.

Second, the two superior types suggest that structural linkages such as teamwork and joint planning should be fostered. Managers should also ensure high market knowledge within marketing. This can be achieved with job rotation policies and joint customer visits of sales and marketing and by obliging executives to serve in the front line from time to time.

Third, the conceptualization provides managers with a systematic way to think through the design of their M & S interface. The multidimensional model of interfaces invites managers to look below the visible surface of information sharing, structural linkages, and power and to manage such "soft factors" as orientations and knowledge. More specifically, managers should work through the following questions: How well do their M & S people know the market? How well do they know the firm's products? and How customer oriented and how product oriented are their M & S people? The most successful cluster manifests particularly high knowledge levels. Through leadership, recruitment, training, and organizational routines, managers have an impact on this intangible resource. Finally, the Brand-Focused Professionals cluster, the most successful configuration, is characterized by a fairly strong marketing function. Thus, the findings challenge the doubts about the necessity of a marketing department that are sometimes uttered in managerial practice (Lorge 1999).

Limitations and Avenues for Further Research

As one of the first empirical studies of the M & S interface, this research has several limitations that should be overcome by future studies. First, further research should conduct an in-depth analysis of structural linkages and integration mechanisms of M & S. This should include analyses of moderating effects—for example, the effect of customer concentration and other context variables not addressed in this article (Rouziès et al. 2005). The finding that the two most successful clusters are characterized by intense structural linkages underscores that this is an important research area. Second, further research should assess how strongly the various interface dimensions discriminate among clusters. Third, further research should collect data from multiple respondents rather than from key informants. Fourth, further research should assume a longitudinal perspective and measure configurations at different points of time. Although this study tested whether configurations Granger-cause profitability, such a data set could also test whether profitability Granger-causes the choice of configuration.

Conclusion

Although the marketing concept subsumes both M & S under the marketing organization, we must accept that M & S frequently are separate functions in managerial practice. The interface between these units has long been neglected in the literature. We hope that this research sheds light on the conceptual dimensions and empirical varieties of this interface and that further research can build on the foundations laid in this article.

APPENDIX
Scale Items for Construct Measures

Construct	Items	Coefficient Alpha/ Composite Reliability
Sales (Versus Marketing) Power •Formative measure based on Homburg, Workman, and Krohmer (1999) •Five-point differential (anchors: 1 = “predominantly marketing’s responsibility,” 3 = “joint responsibility,” and 5 = “predominantly sales’ responsibility”)	Please rate who in your business unit/in your company is primarily responsible for the following tasks: •Communication tasks (e.g., definition of communication activities, design of trade fair appearances) •Market research tasks (e.g., analysis of market potential, planning and execution of a customer satisfaction analysis) •Distribution tasks (e.g., customer relationship management, definition of sales channels) •Service tasks (e.g., definition of product-related services and training offers) •Strategic tasks (e.g., definition of a market strategy) •Product-related tasks (e.g., design and introduction of new products) •Price-related tasks (e.g., definition of price positioning, discounts, and price promotions)	Formative index
Customer (Versus Product) Orientation •Reflective measure based on Cespedes (1995) and Lawrence and Lorsch (1969) •Five-point semantic differential	The sales (marketing) unit of our business unit/company •Aligns volume and revenue plans primarily by … 1 = “products,” and 5 = “customers” •Aligns strategy definition primarily by … 1 = “products,” and 5 = “customers” •Aligns performance evaluations primarily by … 1 = “products,” and 5 = “customers”	Marketing: .85/.88 Sales: .84/.87
Short-Term (Versus Long-Term) Orientation •Reflective measure based on Cespedes (1995) and Lawrence and Lorsch (1969) •Five-point semantic differential	The sales (marketing) unit of our business unit/company is characterized by a … •1 = “systematic/analytical approach,” and 5 = “pragmatic/intuitive approach” The sales (marketing) unit of our business unit/company has a planning horizon that is rather … •1 = “long-term,” and 5 = “short-term”	Marketing .76/.78 Sales .73/.77
Market Knowledge •Formative measure •Five-point scale (anchors: 1 = “low,” and 5 = “high”)	Please think of a typical employee in marketing (sales) in your business unit/in your company. How do you assess this employee with respect to: •Knowledge about customers? •Knowledge about competitors?	Marketing formative index Sales formative index
Product Knowledge •Formative measure •Five-point scale (anchors: 1 = “low,” and 5 = “high”)	Please think of a typical employee in marketing (sales) in your business unit/in your company. How do you assess this employee with respect to: •Knowledge about products? •Knowledge about internal processes?	Marketing formative index Sales formative index
Information Provision •Reflective measure based on Jaworski and Kohli (1993) and Moorman (1995) •Five-point scale (anchors: 1 = “strongly disagree,” and 5 = “strongly agree”)	People in the sales (marketing) unit of our business unit/company … •Are willing to deal with information request from marketing (sales) people. •Respond promptly and without a reminder to information requests from marketing (sales) people. •Inform the marketing (sales) unit proactively.	Marketing .80/.82 Sales .81/.83
Formalization •Reflective measure based on Ruekert and Walker (1987b) •Five-point scale (anchors: 1 = “strongly disagree,” and 5 = “strongly agree”)	To what extent do you agree with the following statements regarding your business unit/company: In the context of M & S cooperation, •We rely on clearly defined coordination processes. •Coordination processes are documented in written form. •We have invested a lot of time into the development of guidelines. •We stick to guidelines (e.g., to rules of cooperation).	.82/.85

APPENDIX
Continued

Construct	Items	Coefficient Alpha/ Composite Reliability
Joint Planning •Reflective measure based on Piercy (1989) •Five-point scale (anchors: 1 = "strongly disagree," and 5 = "strongly agree")	To what extent do you agree with the following statements: In the course of goal definition and planning of M & S within our business unit/company, •Employees from both units are involved. •Market goals and sales goals are reconciled. •Both units commit to common goals. •M&S activities are coordinated with respect to the goals pursued.	.89/.92
Teamwork •Reflective measure based on Cespedes (1996) •Five-point scale (anchors: 1 = "strongly disagree," and 5 = "strongly agree")	To what extent do you agree with the following statements regarding your business unit/company: In the context of M & S cooperation, •We plan many market-related activities jointly in teams (e.g., in project groups, committees). •We take many decisions on market-related activities jointly in teams (e.g., in project groups, committees). •We carry out many market-related activities jointly in teams (e.g., in project groups, committees). •We solve many market-related problems jointly in teams (e.g., in project groups, committees).	.90/.93
Quality of Cooperation Between M & S •Reflective measure based on Ellinger (2000) •Five-point scale (anchors: 1 = "strongly disagree," and 5 = "strongly agree")	To what extent do you agree with the following statements: In our business unit/company, M & S ... •Collaborate frictionless. •Act in concert. •Coordinate their market-related activities. •Have few problems in their cooperation. •Achieve their common goals. •Trust each other.	.91/.92
Market Performance of Business Unit •Reflective measure based on Homburg and Pflesser (2000) •Five-point scale (anchors: 1 = "much worse," 3 = "like," 5 = "much more than competition")	To what extent has your business unit/your company achieved better results than the competition in the following areas: •Achieving customer satisfaction and loyalty •Achieving or maintaining the envisioned market share •Gaining new customers •Making profits •Fast reaction to opportunities and threats in the market	.75/.78
Internal Dynamism •Reflective measure •Five-point scale (anchors: 1 = "very seldom," and 5 = "very frequently")	Please indicate how frequently the following aspects change in your business unit/in your company: •Work processes •Evaluation criteria for employees •Organizational structure •Superiors •Business strategy	.76/.78
Environmental Dynamism •Reflective measure based on Maltz and Kohli (1996) •Five-point scale (anchors: 1 = "very seldom," and 5 = "very frequently")	Please indicate how frequently the following aspects change in the market served by your business unit/by your company: •Products of the competition •Customer needs •Product technology	.64/.69

REFERENCES

Arabie, P. and L. Hubert (1994), "Cluster Analysis in Marketing Research," in *Advanced Methods of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell Business, 160–79.

Armstrong, J. Scott and Terry S. Overton (1977), "Estimating Nonresponse Bias in Mail Surveys," *Journal of Marketing Research*, 14 (August), 396–402.

Bagozzi, Richard P. and Hans Baumgartner (1994), "The Evaluation of Structural Equation Models and Hypothesis Testing," in *Principles of Marketing Research*, Richard P. Bagozzi, ed. Cambridge, MA: Blackwell Business, 386–422.

——— and Youjae Yi (1988), "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 16 (1), 74–94.

Bailey, Kenneth D. (1994), *Typologies and Taxonomies: An Introduction to Classification Techniques*. Thousand Oaks, CA: Sage Publications.

Bourgeois, L.J., III (1985), "Strategic Goals, Environmental Uncertainty, and Economic Performance in Volatile Environments," *Academy of Management Journal*, 28 (3), 548–73.

Brunswik, Egon (1955), "Representative Design and Probabilistic Theory in a Functional Psychology," *Psychological Review*, 62 (3), 193–217.

Bunge, Mario (1996), *Finding Philosophy in Social Science*. New Haven, CT: Yale University Press.

Bunn, Michele D. (1993), "Taxonomy of Buying Decision Approaches," *Journal of Marketing*, 57 (January), 38–56.

Cannon, Joseph P. and William D. Perreault Jr. (1999), "Buyer–Seller Relationships in Business Markets," *Journal of Marketing Research*, 36 (November), 439–60.

Carpenter, Philip (1992), "Bridging the Gap Between Marketing and Sales," *Sales & Marketing Management*, (March), 29–31.

Carper, William B. and William E. Snizek (1980), "The Nature and Types of Organizational Taxonomies: An Overview," *Academy of Management Review*, 5 (1), 65–75.

Cespedes, Frank V. (1993), "Coordinating Sales and Marketing in Consumer Good Firms," *Journal of Consumer Marketing*, 10 (2), 37–55.

——— (1995), *Concurrent Marketing: Integrating Product, Sales, and Service*. Boston: Harvard Business School Press.

——— (1996), "Beyond Teamwork: How the Wise Can Synchronize," *Marketing Management*, 5 (1), 24–37.

Cohen, Jacob, Patricia Cohen, Stephen G. West, and Leona S. Aiken (2003), *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, 3d ed. Mahwah, NJ: Lawrence Erlbaum Associates.

Coviello, Nicole E., Roderick J. Brodie, Peter J. Danaher, and Wesley J. Johnston (2002), "How Firms Relate to Their Markets: An Empirical Examination of Contemporary Marketing Practices," *Journal of Marketing*, 66 (July), 33–46.

———, Heidi Winklhofer, and Karla Hamilton (2006), "Marketing Practices and Performance of Small Service Firms," *Journal of Service Research*, 9 (1), 38–58.

Dahl, Robert A. (1957), "The Concept of Power," *Behavioral Science*, 2 (3), 201–215.

Dastmalchian, Ali and David A. Boag (1990), "Environmental Dependence and Departmental Structure: Case of the Marketing Function," *Human Relations*, 43 (12), 1257–76.

Day, George (1999), "Aligning Organizational Structure to the Market," *Business Strategy Review*, 10 (3), 33–46.

De Ruyter, Ko and Martin Wetzels (2000), "The Marketing-Finance Interface: A Relational Exchange Perspective," *Journal of Business Research*, 50 (2), 209–215.

Deshpandé, Rohit and Frederick E. Webster (1989), "Organizational Culture and Marketing: Defining the Research Agenda," *Journal of Marketing*, 53 (January), 3–15.

Dess, Gregory G., Stephanie Newport, and Abdul M.A. Rasheed (1993), "Configuration Research in Strategic Management: Key Issues and Suggestions," *Journal of Management*, 19 (4), 775–95.

Dewsnap, Belinda and David Jobber (2000), "The Sales-Marketing Interface in Consumer Packaged-Goods Companies: A Conceptual Framework," *Journal of Personal Selling & Sales Management*, 20 (2), 109–119.

——— and ——— (2002), "A Social Psychological Model of Relations Between Marketing and Sales," *European Journal of Marketing*, 36 (7–8), 874–94.

Diamantopoulos, Adamantios and Heidi M. Winklhofer (2001), "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing Research*, 38 (May), 269–77.

Dillon, William R. (2001), "How to Combine Multiple Items into a Composite Score," *Journal of Consumer Psychology*, 10 (1–2), 62.

Donath, Bob (2004), "10 Tips Help Align Sales, Marketing Teams," *Marketing News*, 38 (12), 5–7.

Dooley, Robert S. and Gerald E. Fryxell (1999), "Attaining Decision Quality and Commitment from Dissent: The Moderating Effects of Loyalty and Competence in Strategic Decision-Making Teams," *Academy of Management Journal*, 42 (4), 389–402.

Doty, D. Harold, William H. Glick, and George P. Huber (1993), "Fit, Equifinality, and Organizational Effectiveness: A Test of Two Configurational Theories," *Academy of Management Journal*, 36 (6), 1196–1250.

Dougherty, Deborah (1992), "Interpretive Barriers to Successful Product Innovation in Large Firms," *Organization Science*, 3 (2), 179–202.

Duda, Richard O. and P.E. Hart (1973), *Pattern Classification and Scene Analysis*. New York: John Wiley & Sons.

Dufour, Jean-Marie and Eric Renault (1998), "Short-Run and Long-Run Causality in Time Series: Theory," *Econometrica*, 66 (5), 1099–1125.

Eisenhardt, Kathleen M. and L.J. Bourgeois III (1988), "Politics of Strategic Decision Making in High Velocity Environments: Toward a Midrange Theory," *Academy of Management Journal*, 31 (4), 737–70.

Ellinger, Alexander E. (2000), "Improving Marketing/Logistics Cross-Functional Collaboration in the Supply Chain," *Industrial Marketing Management*, 29 (1), 85–96.

Emerson, Richard M. (1962), "Power-Dependence Relations," *American Sociological Review*, 27 (1), 31–41.

Enz, Cathy A. (1989), "The Measurement of Perceived Intraorganizational Power: A Multi-Respondent Perspective," *Organization Studies*, 10 (2), 241–51.

Fisher, Robert J., Elliot Maltz, and Bernard J. Jaworski (1997), "Enhancing Communication Between Marketing and Engineering: The Moderating Role of Relative Functional Identification," *Journal of Marketing*, 61 (July), 54–70.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39–50.

Frankwick, Gary L., James C. Ward, Michael D. Hutt, and Peter H. Reingen (1994), "Evolving Patterns of Organizational Beliefs in the Formation of Strategy," *Journal of Marketing*, 58 (April), 96–110.

French, John, Jr., and Bertram Raven (1959), "The Basis of Social Power," in *Studies in Social Power*, Dorwin Cartwright, ed. Ann Arbor, MI: Institute for Social Research, 150–67.

Galbraith, Jay R. (1973), *Designing Complex Organizations*. Reading, MA: Addison-Wesley.

Gaski, John F. (1987), "The History of the Measurement of Power in Marketing Channels," in *Review of Marketing*, Michael J. Houston, ed. Chicago: American Marketing Association, 67–89.

——— and John R. Nevin (1985), "The Differential Effects of Exercised and Unexercised Power Sources in a Marketing Channel," *Journal of Marketing Research*, 22 (May), 130–42.

Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37 (3), 424–38.

Griffin, Abbie and John R. Hauser (1996), "Integrating R&D and Marketing: A Review and Analysis of the Literature," *Journal of Product Innovation Management*, 13 (3), 191–215.

Harzing, Anne-Will (1997), "Response Rates in International Mail Surveys: Results of a 22-Country Study," *International Business Review*, 6 (6), 641–64.

Helsen, Kristiaan and Paul E. Green (1991), "A Computational Study of Replicated Clustering with an Application to Market Segmentation," *Decision Sciences*, 22 (5), 1124–41.

Hinings, C.R., David J. Hickson, Johannes M. Pennings, and R.E. Schneck (1974), "Structural Conditions of Intraorganizational Power," *Administrative Science Quarterly*, 19 (1), 22–44.

Homburg, Christian and Christian Pflesser (2000), "A Multiple-Layer Model of Market-Oriented Organizational Culture: Measurement Issues and Performance Outcomes," *Journal of Marketing Research*, 37 (November), 449–62.

———, John P. Workman Jr., and Ove Jensen (2000), "Fundamental Changes in Marketing Organization: The Movement Toward a Customer-Focused Organizational Structure," *Journal of the Academy of Marketing Science*, 28 (4), 459–79.

———, ———, and ——— (2002), "A Configurational Perspective on Key Account Management," *Journal of Marketing*, 66 (April), 38–60.

———, ———, and Harley Krohmer (1999), "Marketing's Influence Within the Firm," *Journal of Marketing*, 63 (April), 1–17.

Hunt, Shelby D. and John R. Nevin (1974), "Power in a Channel of Distribution: Sources and Consequences," *Journal of Marketing Research*, 11 (May), 186–93.

Jarvis, Cheryl B., Scott B. Mackenzie, and Philip M. Podsakoff (2003), "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research*, 30 (September), 199–218.

Jaworski, Bernard J. and Ajay K. Kohli (1993), "Market Orientation: Antecedents and Consequences," *Journal of Marketing*, 57 (July), 53–70.

Kahn, Kenneth B. (1996), "Interdepartmental Integration: A Definition with Implications for Product Development Performance," *Journal of Product Innovation Management*, 13 (2), 137–51.

———, Richard C. Reizenstein, and Joseph O. Rentz (2004), "Sales-Distribution Interfunctional Climate and Relationship Effectiveness," *Journal of Business Research*, 57 (10), 1085–1091.

Ketchen, David J., James G. Combs, Craig J. Russell, Chris Shook, Michelle A. Dean, Janet Runge, et al. (1997), "Organizational Configurations and Performance: A Meta-Analysis," *Academy of Management Journal*, 40 (1), 223–40.

——— and Christopher L. Shook (1996), "The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique," *Strategic Management Journal*, 17 (6), 441–58.

———, James B. Thomas, and Charles C. Snow (1993), "Organizational Configurations and Performance: A Comparison of Theoretical Approaches," *Academy of Management Journal*, 36 (6), 1278–1313.

Kirca, Ahmet H., Satish Jayachandran, and William O. Bearden (2005), "Market Orientation: A Meta-Analytic Review and Assessment of Its Antecedents and Impact on Performance," *Journal of Marketing*, 69 (April), 24–41.

Kotler, Philip, Neil Rackham, and Suj Krishnaswamy (2006), "Ending the War Between Sales & Marketing," *Harvard Business Review*, 84 (7–8), 68–78.

Krohmer, Harley, Christian Homburg, and John P. Workman Jr. (2002), "Should Marketing Be Cross-Functional?" *Journal of Business Research*, 55 (6), 451–65.

Lawrence, Paul R. and Jay W. Lorsch (1969), *Organization and Environment: Managing Differentiation and Integration*. Homewood, IL: Richard D. Irwin.

Leenders, Mark A. and Berend Wierenga (2002), "The Effectiveness of Different Mechanisms for Integrating Marketing and R&D," *Journal of Product Innovation Management*, 19 (4), 305–317.

Lorge, Sarah (1999), "Marketers Are From Mars, Salespeople Are From Venus," *Sales & Marketing Management*, 151 (4), 27–33.

Lusch, Robert F. and James R. Brown (1982), "A Modified Model of Power in the Marketing Channel," *Journal of Marketing Research*, 19 (August), 312–23.

Maltz, Elliot and Ajay K. Kohli (1996), "Market Intelligence Dissemination Across Functional Boundaries," *Journal of Marketing Research*, 33 (February), 47–61.

——— and ——— (2000), "Reducing Marketing's Conflict with Other Functions: The Differential Effects of Integrating Mechanisms," *Journal of the Academy of Marketing Science*, 28 (4), 479–92.

———, William E. Souder, and Ajith Kumar (2001), "Influencing R&D/Marketing Integration and the Use of Market Information R&D Managers: Intended and Unintended Effects of Managerial Actions," *Journal of Business Research*, 52 (1), 69–82.

Marinova, Detelina (2004), "Actualizing Innovation Effort: The Impact of Market Knowledge Diffusion in a Dynamic System of Competition," *Journal of Marketing*, 68 (July), 1–20.

Martinez, Jon I. and J. Carlos Jarillo (1989), "The Evolution of Research on Coordination Mechanisms in Multinational Corporations," *Journal of International Business Studies*, 20 (3), 489–514.

Mason, Richard O. and Ian I. Mitroff (1981), *Challenging Strategic Planning Assumptions*. New York: John Wiley & Sons.

McIntyre, Robert M. and Roger K. Blashfield (1980), "A Nearest-Centroid Technique for Evaluating the Minimum-Variance Clustering Procedure," *Multivariate Behavioral Research*, 15 (2), 225–38.

McKelvey, Bill (1975), "Guidance for the Empirical Classification of Organizations," *Administrative Science Quarterly*, 20 (4), 509–525.

Merton, Robert K. (1968), *Social Theory and Social Structure*. New York: The Free Press.

Meyer, Alan D., Anne S. Tsui, and C.R. Hinings (1993), "Configurational Approaches to Organizational Analysis," *Academy of Management Journal*, 36 (6), 1175–95.

Miles, Raymond E. and Charles C. Snow (1978), *Organizational Strategy, Structure, and Process*. New York: McGraw-Hill.

Milligan, Glenn W. (1996), "Clustering Validation: Results and Implications for Applied Analyses," in *Clustering and Classification*, P. Arabie, L. Hubert, and G. De Soete, eds. River Edge, NJ. World Scientific Publishing, 341–75.

——— and Martha C. Cooper (1985), "An Examination of Procedures for Determining the Number of Clusters in a Data Set," *Psychometrika*, 50 (2), 129–79.

——— and ——— (1987), "Methodology Review: Clustering Methods," *Applied Psychological Measurement*, 11 (4), 329–54.

——— and S.C. Hirtle (2003), "Clustering and Classification Methods," in *Handbook of Psychology: Research Methods in*

Psychology, Vol. 2, J.A. Schinka, W.F. Velicer, and I.B. Weiner, eds. New York: John Wiley & Sons, 165–86.

Mitroff, Ian I. (1974), "A Brunswik Lens Model of Dialectical Inquiring Systems," *Theory and Decision*, 5 (1), 45–67.

Montgomery, David B. and Frederick E. Webster Jr. (1997), "Marketing's Interfunctional Interfaces: The MSI Workshop on Management of Corporate Fault Zones," *Journal of Market-Focused Management*, 2 (1), 7–26.

Moorman, Christine (1995), "Organizational Market Information Processes: Cultural Antecedents and New Product Outcomes," *Journal of Marketing Research*, 32 (August), 318–35.

——— and Roland T. Rust (1999), "The Role of Marketing," *Journal of Marketing*, 63 (October), 180–97.

Olson, Eric M., Stanley F. Slater, and G. Tomas M. Hult (2005), "The Performance Implications of Fit Among Business Strategy, Marketing Organization Structure, and Strategic Behavior," *Journal of Marketing*, 69 (July), 49–65

Panigyrakis, George and Cleopatra Veloutsou (1999), "Brand Managers' Interfaces in Different Consumer Goods Industries," *Journal of Product & Brand Management*, 8 (1), 19–37.

Piercy, Nigel (1986), "The Role and Function of the Chief Marketing Executive and the Marketing Department: A Study of Medium-Sized Companies in the UK," *Journal of Marketing Management*, 1 (3), 265–89.

——— (1989), "The Power and Politics of Sales Forecasting: Uncertainty Absorption and the Power of the Marketing Department," *Journal of the Academy of Marketing Science*, 17 (2), 109–120.

Pinder, Craig C. and Larry F. Moore (1979), "The Resurrection of Taxonomy to Aid the Development of Middle Range Theories of Organizational Behavior," *Administrative Science Quarterly*, 24 (1), 99–118.

Podsakoff, Philip M., Scott B. MacKenzie, Lee Jeong-Yeon, and Nathan P. Podsakoff (2003), "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, 88 (5), 879–903.

Punj, Girish and David W. Stewart (1983), "Cluster Analysis in Marketing Research: Review and Suggestions for Application," *Journal of Marketing Research*, 20 (May), 134–48.

Rand, William M. (1971), "Objective Criteria for the Evaluation of Clustering Methods," *Journal of the American Statistical Association*, 66 (336), 846–50.

Reilly, Anne H., Jeanne M. Brett, and Linda K. Stroh (1993), "The Impact of Corporate Turbulence on Employee Attitudes," *Strategic Management Journal*, 14 (4), 167–79.

Rich, Philip (1992), "The Organizational Taxonomy: Definition and Design," *Academy of Management Review*, 17 (4), 758–81.

Rouziès, Dominique, Erin Anderson, Ajay K. Kohli, Ronald E. Michaels, Barton A. Weitz, and Andris A. Zoltners (2005), "Sales and Marketing Integration: A Proposed Framework," *Journal of Personal Selling & Sales Management*, 25 (2), 113–22.

Ruekert, Robert W. and Orville C. Walker Jr. (1987a), "Interactions Between Marketing and R&D Departments in Implementing Different Business Strategies," *Strategic Management Journal*, 8 (3), 233–48.

——— and ——— (1987b), "Marketing's Interaction with Other Functional Units: A Conceptual Framework and Empirical Evidence," *Journal of Marketing*, 51 (January), 1–19.

———, ———, and Kenneth J. Roering (1985), "The Organization of Marketing Activities: A Contingency Theory of Structure and Performance," *Journal of Marketing*, 49 (January), 13–25.

Sarle, W.S. (1983), *Cubic Clustering Criterion*. Cary, NC: SAS Institute Inc.

Scott, W. Richard (1992), *Organizations: Rational, Natural and Open Systems*. Englewood Cliffs, NJ: Prentice Hall.

Slater, Stanley F. and Eric M. Olson (2001), "Marketing's Contribution to the Implementation of Business Strategy: An Empirical Analysis," *Strategic Management Journal*, 22 (11), 1055–1067.

Smith, Timothy M., Srinath Gopalakrishna, and Rabikar Chatterjee (2006), "A Three-Stage Model of Integrated Marketing Communications at the Marketing–Sales Interface," *Journal of Marketing Research*, 43 (November), 564–79.

Sneath, Peter H.A. and Robert R. Sokal (1973), *Numerical Taxonomy*. San Francisco: Freeman.

Song, X. Michael, Jinhong Xie, and Barbara Dyer (2000), "Antecedents and Consequences of Marketing Managers' Conflict-Handling Behaviors," *Journal of Marketing*, 64 (January), 50–66.

Starr, Richard G., Jr., and Paul N. Bloom (1994), "The Power Relationships of Brand Managers," *Marketing Letters*, 5 (3), 211–23.

Strahle, William M., Rosann L. Spiro, and Frank Acito (1996), "Marketing and Sales: Strategic Alignment and Functional Implementation," *Journal of Personal Selling & Sales Management*, 16 (1), 1–20.

Tushman, Michael L. and Elaine Romanelli (1994), "Organizational Transformation as Punctuated Equilibrium: An Empirical Test," *Academy of Management Journal*, 37 (5), 1141–66.

Venkatraman, N. (1989), "The Concept of Fit in Strategy Research: Toward Verbal and Statistical Correspondence," *Academy of Management Review*, 14 (3), 423–44.

Vorhies, Douglas W. and Neil A. Morgan (2003), "A Configuration Theory Assessment of Marketing Organization Fit with Business Strategy and Its Relationship with Marketing Performance," *Journal of Marketing*, 67 (January), 100–115.

Waller, Ray A. and David B. Duncan (1969), "A Bayes Rule for the Symmetric Multiple Comparisons Problem," *Journal of the American Statistical Association*, 64 (December), 1484–1503.

Ward, J.H., Jr. (1963), "Hierarchical Grouping to Optimize an Objective Function," *Journal of the American Statistical Association*, 58 (301), 236–44.

Webster, Frederick E., Jr. (1997), "The Future Role of Marketing in the Organization," in *Reflections on the Futures of Marketing*, Donald R. Lehmann and Katherine E. Jocz, eds. Cambridge, MA: Marketing Science Institute, 39–66.

Weitz, B.A. and E. Anderson (1981), "Organizing the Marketing Function," in *Review of Marketing*, B. Enis and K. Roering, eds. Chicago: American Marketing Association, 134–42.

Workman, John P., Jr. (1993), "Marketing's Limited Role in New Product Development in One Computer Systems Firm," *Journal of Marketing Research*, 30 (November), 405–421.

———, Christian Homburg, and Kjell Gruner (1998), "Marketing Organization: An Integrative Framework of Dimensions and Determinants," *Journal of Marketing*, 62 (July), 21–41.

Check out our other AMA publications

Call now to learn how AMA magazines can help prepare you and your company for the challenges ahead.

TO SUBSCRIBE call 800-262-1150 or go to www.amaorders.com

Marketing Research

Marketing Research is tailored specifically to managers of marketing research activities. Researchers and managers count on this quarterly resource to help stay on top of current methodologies and issues, management concerns, and the latest books and oftware. Inside you'll find in-depth analysis of research applications and heated debates that challenge conventional ideas about the state of marketing research today—and tomorrow.

Marketing Management

Designed to serve busy executives, *Marketing Management* focuses on strategic marketing issues that marketing managers face every day. Our pages are packed with expert insights from today's thought leaders as well as case studies and interviews with marketing executives. Published six times a year, *Marketing Management* sheds light on hot topics—like brand management, CRM, marketing technology, global marketing, B2B, and services marketing—to help managers keep pace with this rapidly changing field.

Marketing Health Services

Marketing professionals looking for new ways to market their healthcare organization need look no further. *Marketing Health Services*, a quarterly magazine specifically aimed at senior-level healthcare marketers and managers, offers targeted information, practical strategies, and thought-provoking commentary to help achieve your goals and shape your vision. Regular topics include e-health, DTC marketing, legislative developments, database marketing, and more.

MARKETING

NEW MARKETING TEXTBOOK

Don't make your students choose between:

- New textbook, price $150 or more
- Second hand textbook

*Assign them a **great textbook** at an **affordable** price*

$45 hard copy / **$25** electronic

Managing Marketing in the 21st Century

Examine ***Managing Marketing in the 21st Century*** online at **www.mm21c.com** (instructors' area available)

Noel Capon (author of *The Marketing Mavens*)
with
James Mac Hulbert
Columbia Business School

wessex

www.mm21c.com

Journal of
Marketing

Vol. 72, No. 3

CENTRAL LIBRARY
1.7.08
UNIVERSITY
OF
CALCUTTA
ALIPORE CAMPUS

AMERICAN
MARKETING
ASSOCIATION

EDITOR
Roland T. Rust
Robert H. Smith School of Business
3451 Van Munching Hall
University of Maryland
College Park, MD 20742-1815

EDITOR-ELECT
Ajay Kohli
Emory University

ASSISTANTS TO THE EDITOR
Chris Bartone
E-mail: jom@ama.org

Millicent Locke
University of Maryland

WEB EDITOR
Siva K. Balasubramanian
Southern Illinois University

BOOK REVIEWS
Ramarao Desiraju
University of Central Florida

VICE PRESIDENT—PUBLICATIONS
Russell S. Winer
New York University

CHIEF EXECUTIVE OFFICER
Dennis Dunlap

GROUP PUBLISHER
Jack Hollfelder

EDITORIAL STAFF
Phone: (312) 542-9000
Fax: (312) 542-9001

MANAGING EDITOR
Francesca Van Gorp Cooley

TECHNICAL EDITOR
Andy Seagram

GRAPHICS EDITOR
Mary Loye

Advertising Staff
Fax: (312) 922-3763

Richard Ballschmiede, Advertising
(312) 542-9076

Sally Schmitz, Production
(312) 542-9038

JM Web site:
Submissions: www.marketingjournals.org/JM
Main page: www.marketingpower.com/JM

The *Journal of Marketing* (ISSN 0022-2429) is published bimonthly (January, March, May, July, September, and November). *Annual Subscription Rates*: Subscriptions are $115 for individuals in the United States and Canada and $160 in other countries. Library, school, and business subscriptions are $325 in the United States and Canada and $360 in other countries. Subscriptions for AMA members are $60. Prices are for one year. Single copy rate: $50 individual, $60 corporate in the United States and Canada, add $5 in other countries Prepayment is required in U.S. dollars or the equivalent.

Published by the Publications Group of the American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266. Periodicals postage paid at Chicago, IL and additional mailing offices. POSTMASTER. Send address changes to the *Journal of Marketing*, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266, USA.

•**MANUSCRIPTS**. Electronic submissions must be made via the *JM* Web site at http://www.marketingjournals.org. Paper submissions are not accepted. Editorial policies, review board members, Tables of Contents for *JM*, and so on, are available on the *JM* Web site at http://www.marketingpower.com/jm.

•**ADVERTISERS AND ADVERTISING AGENCIES** assume liability for all content (including text, representations, and illustrations) of advertisements published, and also assume responsibility for any claims arising therefrom made against the publisher The right is reserved to reject any advertisement.

•**SUBSCRIPTIONS:** Members· Send to American Marketing Association, 311 S. Wacker Dr., Suite 5800, Chicago, IL 60606-2266 (1-800-AMA-1150 or 312/542-9000), or e-mail to info@ama.org. Nonmembers: Send to American Marketing Association, P.O. Box 11806, Birmingham, AL 35202, 1-800-633-4931, 205-995-1567, Fax: 205-995-1588, e-mail. amasubs@ebsco.com. Canadian residents add 6% GST (GST registration # 12747852).

•**PERMISSIONS AND REPRINT INFORMATION:** Without written permission from the AMA, any copying or reprinting (except by authors reprinting their own work) is prohibited. Requests for permission to reprint—such as copying for general distribution, advertising, or promotional purpose; for creating new collective works, or for resale—should be submitted in writing to permissions@ama.org. Bulk reprints are available by contacting Tamara Groft, Sheridan Reprints, (800) 352-2210. Permission is granted by AMA for those registered with the Copyright Clearance Center (CCC), 222 Rosewood Dr., Danvers, MA 01923, to photocopy any article herein for a flat fee Contact CCC with ISSN 0022-2429, volume, and first and last page numbers of each article copied.

The views expressed in *JM* are not necessarily those of the editorial staff or the publisher. **All articles in this journal, unless otherwise noted, have undergone a blind review process.**

Copyright © 2008, American Marketing Association. All rights reserved.

Editorial Review Board

ROHINI AHLUWALIA
University of Minnesota

KUSUM L AILAWADI
Dartmouth College

JOSEPH ALBA
University of Florida

TIM AMBLER
London Business School

EUGENE W. ANDERSON
University of Michigan

TOR WALLIN ANDREASSEN
Norwegian School of Management

ASIM ANSARI
Columbia University

ERIC J. ARNOULD
University of Wyoming

KWAKU ATUAHENE-GIMA
CEIBS

BARRY L BAYUS
University of North Carolina

WILLIAM BEARDEN
University of South Carolina

PAUL D. BERGER
Boston University

PAUL N. BLOOM
Duke University

LISA E. BOLTON
University of Pennsylvania

RUTH N BOLTON
Arizona State University

ROBERT BORDLEY
General Motors Corporation

DOUGLAS BOWMAN
Emory University

RODERICK J BRODIE
University of Auckland

STEVEN P. BROWN
University of Houston

JOSEPH CANNON
Colorado State University

DIPANKAR CHAKRAVARTI
University of Colorado

PIERRE CHANDON
INSEAD

YUXIN CHEN
New York University

ALEXANDER CHERNEV
Northwestern University

PRADEEP K. CHINTAGUNTA
University of Chicago

KEVIN CLANCY
Copernicus Marketing Consulting

PETER J DANAHER
University of Melbourne

KO DE RUYTER
Maastricht University

JOHN A DEIGHTON
Harvard University

MARNIK G DEKIMPE
Tilburg University and Catholic University Leuven

WAYNE S. DESARBO
Pennsylvania State University

RAVI DHAR
Yale University

NAVEEN DONTHU
Georgia State University

TÜLIN ERDEM
New York University

PETER S FADER
University of Pennsylvania

GARY L FRAZIER
University of Southern California

GARY F GEBHARDT
University of South Florida

INGE GEYSKENS
Tilburg University

PETER N. GOLDER
New York University

RAJDEEP GREWAL
Pennsylvania State University

ABBIE GRIFFIN
University of Utah

SUNIL GUPTA
Harvard University

ZEYNEP GÜRHAN-CANLI
KOC University

MICHAEL HAENLEIN
ESCP-EAP

JOHN R HAUSER
Massachusetts Institute of Technology

JAN B. HEIDE
University of Wisconsin–Madison

CHRISTIAN HOMBURG
University of Mannheim

WAYNE D. HOYER
University of Texas at Austin

JOEL HUBER
Duke University

JOHN HULLAND
University of Pittsburgh

G. TOMAS M. HULT
Michigan State University

J. JEFFREY INMAN
University of Pittsburgh

GANESH IYER
University of California, Berkeley

SANJAY JAIN
Texas A&M University

CHRIS JANISZEWSKI
University of Florida

SANDY D. JAP
Emory University

BERNARD J. JAWORSKI
Monitor Group

P.K. KANNAN
University of Maryland

TIMOTHY L KEININGHAM
Ipsos Loyalty

AMNA KIRMANI
University of Maryland

AJAY K. KOHLI
Emory University

ROBERT V. KOZINETS
York University

V. KUMAR
University of Connecticut

DONALD R LEHMANN
Columbia University

KATHERINE N. LEMON
Boston College

MARY FRANCES LUCE
Duke University

XUEMING LUO
University of Texas at Arlington

ROBERT F LUSCH
University of Arizona

JOHN G LYNCH JR
Duke University

VIJAY MAHAJAN
University of Texas at Austin

RICHARD G MCFARLAND
Kansas State University

CARL F. MELA
Duke University

VIKAS MITTAL
University of Pittsburgh

WENDY MOE
University of Maryland

CHRISTINE MOORMAN
Duke University

JOHN MOWEN
Oklahoma State University

EITAN MULLER
New York University and Tel-Aviv University

SCOTT A NESLIN
Dartmouth College

RICHARD NETEMEYER
University of Virginia

RAKESH NIRAJ
University of Southern California

STEPHEN M NOWLIS
Arizona State University

JOSEPH C NUNES
University of Southern California

A. PARASURAMAN
University of Miami

KOEN PAUWELS
Dartmouth College

WILLIAM D PERREAULT
University of North Carolina

ROBERT PETERSON
University of Texas at Austin

F G.M. PIETERS
Tilburg University

JAIDEEP PRABHU
Imperial College

WILLIAM P PUTSIS
University of North Carolina

ARVIND RANGASWAMY
Pennsylvania State University

AKSHAY RAO
University of Minnesota

VITHALA RAO
Cornell University

BRIAN T RATCHFORD
University of Texas at Dallas

WERNER REINARTZ
INSEAD

ARIC RINDFLEISCH
University of Wisconsin–Madison

FRED SELNES
Norwegian School of Management

STEVEN M. SHUGAN
University of Florida

ITAMAR SIMONSON
Stanford University

STANLEY F SLATER
Colorado State University

REBECCA SLOTEGRAAF
Indiana University

JOYDEEP SRIVASTAVA
University of Maryland

JAN-BENEDICT E.M STEENKAMP
University of North Carolina

DAVID W STEWART
University of Southern California

STEFAN STREMERSCH
Erasmus University Rotterdam

BAOHUNG SUN
Carnegie Mellon University

GERARD J. TELLIS
University of Southern California

JACQUELYN THOMAS
Northwestern University

CRAIG THOMPSON
University of Wisconsin–Madison

DEBORA THOMPSON
Georgetown University

CHRISTOPHE VAN DEN BULTE
University of Pennsylvania

STIJN VAN OSSELAER
Erasmus University Rotterdam

P RAJAN VARADARAJAN
Texas A&M University

STEPHEN L VARGO
University of Hawaii, Manoa

SAJEEV VARKI
University of South Florida

RAJKUMAR VENKATESAN
University of Virginia

PETER VERHOEF
University of Groningen

KENNETH WATHNE
University of Wisconsin–Madison

MICHEL WEDEL
University of Maryland

BARTON A WEITZ
University of Florida

RUSSELL S WINER
New York University

CAROLYN YOON
University of Michigan

ANTHONY ZAHORIK
Burke Institute

VALARIE A ZEITHAML
University of North Carolina, Chapel Hill

JIE ZHANG
University of Maryland

Z JOHN ZHANG
University of Pennsylvania

Joseph Johnson & Gerard J. Tellis

Drivers of Success for Market Entry into China and India

China and India are the fastest-growing major markets in the world and the most popular markets for foreign entrants. However, no study has examined the success or failure of these entries. Using a new definition of success and a uniquely compiled archival database, the authors analyze whether and why firms that entered China and India succeeded or failed. The most important findings are rather counterintuitive: Smaller firms are more successful than larger firms, and firms entering more open emerging markets have less success. Other findings are that success is greater with earlier entry, greater control of entry mode, and shorter cultural and economic distances between the home and the host countries. Importantly, with or without control for these drivers, firms have less success in India than in China. The authors discuss the reasons for and implications of these findings.

Keywords: market entry, India, China, success factors

China and India have become major players in the world economy. For example, China and India have led all world economies with gross domestic product (GDP) growth rates of more than 9% in recent years (Vietor 2007). Because of this rapid growth, China and India are currently the third- and fifth-largest economies in purchasing power parity (Wilson and Purushothaman 2003). Some forecasts suggests that by 2020, China and India will pass Japan in GDP in purchasing power parity and that by 2050 China will be the leading economy of the world, followed by the United States and India (Hawksworth 2006).

This remarkable economic resurgence and future promise of China and India have made entering these markets critical to the survival and success of many firms (Wilson and Purushothaman 2003). Of the *Fortune* 500, 400 firms now operate in China (Fishman 2005), and 220 of the top 500 firms operate in India (India Brand Equity Foundation 2005). In 2005, China alone attracted approximately $1 billion per week in foreign direct investment. Whereas firms in the earlier years rushed into these countries primarily for reasons such as acquiring resources, securing key supplies, accessing low-cost factors, and diversifying sources of supply (Vernon, Wells, and Rangan 1996), the rising income of the local populace is now resulting in market-seeking behavior.

How have foreign entrants performed in these emerging markets? What drivers have led to their success or failure? Firms have been reluctant to divulge specific information on performance, and researchers have neglected to study this issue; thus, it has gone largely unexamined. As a result, despite almost three decades of history, it is unclear how firms should enter such emerging markets. Examples of unexplained success and failure abound. Unilever launched 14 joint ventures in China from 1986 to 1999 (Dasgupta and Dutta 2004) and was in the red for most of the time. On the contrary, Procter & Gamble (P&G) ended up as the market leader in almost all categories it introduced in China (Tunistra 2000).

Although the few empirical studies on entry success (e.g., Gielens and Dekimpe 2007; Luo 1998; Pan, Li, and Tse 1999) have made important contributions to the topic, they suffer from at least one of the following limitations: First, the studies focus on a single country—China in most cases. Second, the studies use a restrictive definition of success, such as market share, which does not encapsulate degrees of success and failure. Third, the studies often focus on one particular industry. Fourth, the studies do not cover success or failure over time from the beginning of the liberalization of the Chinese and Indian economies. Against this setting, it is unclear whether these findings are generalizable across industries and emerging markets.

The current study attempts to analyze the success and failure of firms entering the major emerging markets of China and India. It addresses the following research questions: What drives the success of entry into China and India? Is entry into China more or less successful than that into India? and How do entry mode, entry timing, and firm size (firm-level variables) and economic distance, cultural distance, country risk, and country openness (country-level variables) affect success?

Our contributions to the literature are as follows: First, we propose a richer measure for success and failure, which encapsulates longitudinal historical accounts. Second, we relate our measure of success to underlying drivers that emerge from a vast body of interdisciplinary research over decades. Third, we focus on both the major emerging markets: China and India. Fourth, because of the paucity of systematic or syndicated data, we use the historical method (Golder and Tellis 1993) to collect data to answer these questions.

Joseph Johnson is Assistant Professor of Marketing, School of Business Administration, University of Miami (e-mail: jjohnson@miami.edu). Gerard J. Tellis is Director of Center for Global Innovation, Neely Chair of American Enterprise, and Professor of Marketing, Marshall School of Business, University of Southern California, Los Angeles (e-mail: tellis@usc.edu). This article benefited by a grant from Don Murray to the USC Marshall Center for Global Innovation.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

We organize the rest of the article into three sections. First, we discuss the drivers of success or failure and pose specific research questions. Second, we describe the method and results of historical analysis. Third, we discuss the findings, implications, and limitations of our study.

The Drivers of Entry Success

Researchers have not yet developed a single coherent theory of the drivers of success or failure of entry in emerging markets. This section reviews the prior literature on international market entry to identify the drivers of success or failure in market entry, proposes a conceptual framework for these drivers, and derives some questions for empirical research. The interdisciplinary literature spans marketing, strategy, and international business (Dunning 1988; Root and Ahmed 1979; Zhao, Luo, and Suh 2004). We use the terms "firm" to describe the entrant, "host country" to describe China or India, "home country" to describe the firm's country of origin, and "foreign country" to describe any other country that may be involved.

We suggest two broad constructs that drive firm performance in international entry: firm differentiation and country differentiation. Within firm differentiation, two key constructs are firm strategy and firm resources. The most important strategies in international entry are entry mode and entry timing. We measure firm resources with one key variable, firm size.

Within country differentiation, the key construct is host-country characteristics. Among these characteristics, the two that we identify as important are country openness and country risk.

In addition to these constructs, firm and country differentiation together shape host–home location. Two variables of this latter construct that are most extensively discussed in the literature are cultural distance and economic distance. We measure firm performance by the historical success of firms as reported in archival records.

Figure 1 shows a conceptual framework of how the constructs are related to one another and which variables we use to measure these constructs. The subsequent subsections discuss the role of each of these independent variables in affecting historical success or failure.

Entry Mode

The mode of entry is a fundamental decision a firm makes when it enters a new market because the choice of entry automatically constrains the firm's marketing and production strategy. The mode of entry also affects how a firm faces the challenges of entering a new country and deploying new skills to market its product successfully (Gillespie, Jeannet, and Hennessy 2007). A firm entering a foreign market faces an array of choices to serve the market. In an exhaustive survey of the different modes of market entry, Root (1994) identifies 15 different forms. Following Root, we categorize these into the following five main classes, listed in order of increasing control:

1. *Export*: a firm's sales of goods/services produced in the home market and sold in the host country through an entity in the host country.
2. *License and franchise*: a formal permission or right offered to a firm or agent located in a host country to use a home firm's proprietary technology or other knowledge resources in return for payment.
3. *Alliance*: agreement and collaboration between a firm in the home market and a firm located in a host country to share activities in the host country.
4. *Joint venture*: shared ownership of an entity located in a host country by two partners, one located in the home country and the other located in the host country.
5. *Wholly owned subsidiary*: complete ownership of an entity located in a host country by a firm located in the home country to manufacture or perform value addition or sell goods/services in the host country.

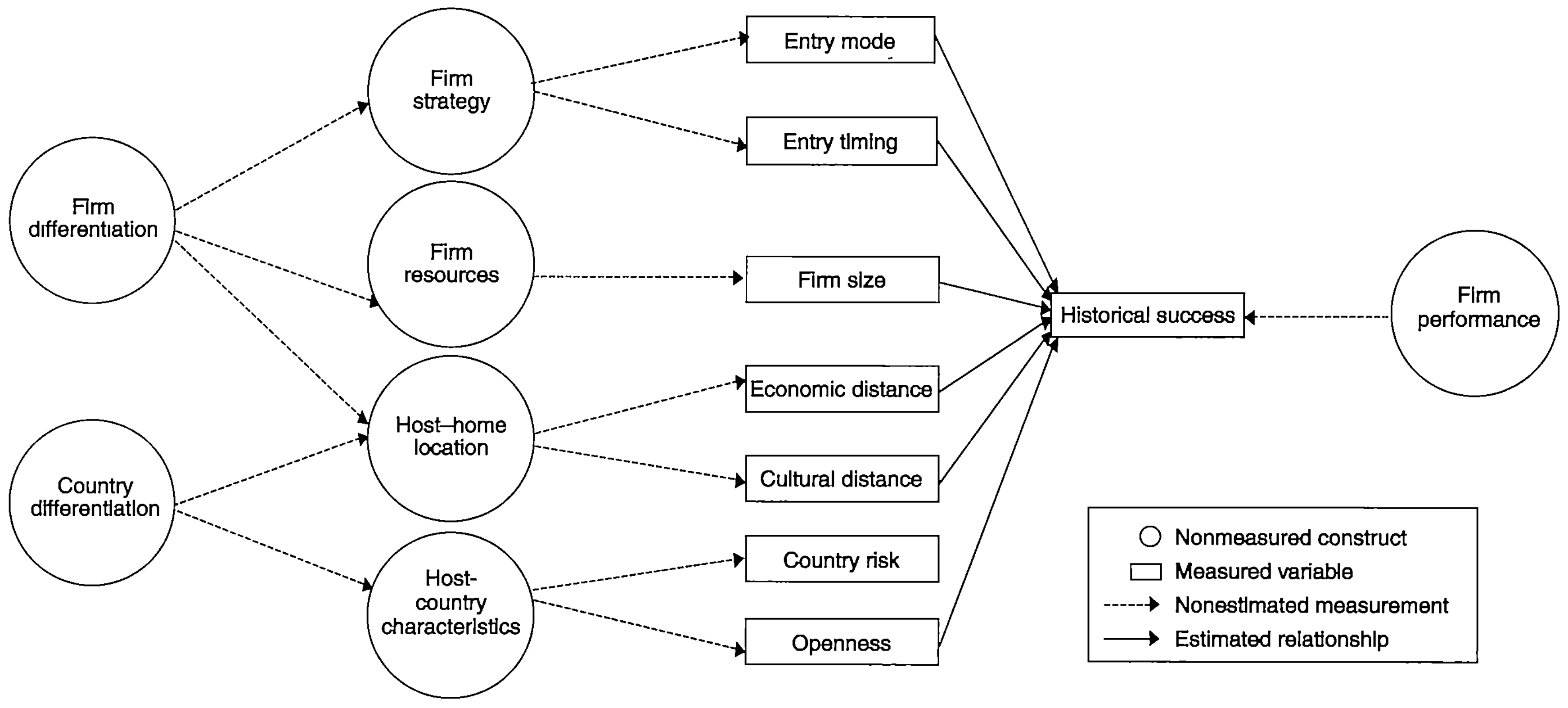
A firm can choose any of these entry modes or some combination of them to enter a host country. The key attribute that distinguishes the different modes of entry is the degree to which they give a firm control over its key marketing resources (Anderson and Gatignon 1986). At one end of the spectrum is the export of goods, which has the lowest degree of control. Licenses, franchises, and various forms of joint venture provide a progressively increasing degree of control for the firm; at the other end of the spectrum, ownership-based entries, such as wholly owned subsidiaries, afford the highest control.

Two opposing theories suggest alternative outcomes as control increases: the resource-based view and the transactions cost view. The resource-based view holds that as the degree of control increases, the firm's chances of success increase because the firm can deploy key resources that are essential to success (Gatignon and Anderson 1988; Isobe, Makino, and Montgomery 2000). These resources can be intangible properties, such as brand equity and marketing knowledge (Arnold 2004), or tangible properties, such as a patent or a process blueprint. Control over such properties gives a firm the freedom to deploy resources flexibly, thus enhancing its chances of success. In the context of emerging markets, control provides two key benefits. First, it safeguards key resources from leakage, such as patent theft. Second, it allows for internal operational control, which is essential to a firm's success in emerging markets (Luo 2001). In addition, a firm can control key complementary resources, such as access to local distribution channels, which can be important to its success in any country.

The transaction cost view holds that costs increase with increasing control of the mode of entry. Control and commitment are inextricably linked in mode of entry (Luo 2001). High control in entry strategies entails high commitment. Transaction cost theory suggests that the higher the resource commitment and desired control of an entry mode, the higher is the cost. Wholly owned subsidiaries and joint ventures are high-cost entry modes because of the level of resource commitment needed to set up operations (Pan and Chi 1999). These higher costs imply that higher levels of investments are needed for the firm to break even and make a profit. Taken together, these arguments lead to our first specific research question:

> Q_1: Does success in entering emerging markets increase or decrease with the degree of control?

FIGURE 1
Conceptual Framework: Drivers of Entry Success



Entry Timing

In addition to the entry mode, the role of market entry timing is critical in emerging markets (Pan and Chi 1999). However, the direction of the effect is not clear. The literature suggests reasons that early entry into international markets could favor or hurt success.

On the one hand, early entry has many advantages. First, the early entrant can lock up access to key resources, such as distribution channels and suppliers. Second, early entrants have the opportunity to set the pattern of consumer preference (Carpenter and Nakamoto 1989; Mitchell 1999), which may disadvantage later entrants. Third, early entrants can benefit from being the first to exploit governmental concessions and incentives, which governments often offer to attract such entrants (Pan and Chi 1999). Fourth, early entrants can time their entry to exploit the "strategic window" of an expanding market and observe and learn market attributes for a longer period. Pan and Chi (1999, p. 360) report that "[multinational corporations] that started their production in China in an earlier year had a higher level of profit than those that began in a later year."

On the other hand, Golder and Tellis (1993) find that pioneers are often not the long-term winners in a market. Using U.S. data, they show that in several categories, "best" beats "first" (Tellis and Golder 2001). In the international context, pioneers may fail for several reasons. First, firms that rush in first may not be aware of the pitfalls of the newly opened emerging market. Second, returns to the early entrants might be too low compared with their investments, especially because infrastructure is not yet fully developed. Third, latter entrants have a flatter learning curve because they can learn from the early entrants' errors (Fujikawa and Quelch 1998). These three factors may be responsible for the failure of many early entrants in some markets (Arnold 2004). These arguments lead to our second research question:

> Q_2: Does success in entering emerging markets increase or decrease with early entry?

Firm Size

New trade theories developed by Krugman (1980) and Porter (1990) suggest that firm-specific advantages play an important role in international trade. Although small firms (with fewer than 500 employees) today account for 30% of U.S. exports (Cateora and Graham 2006), in general, larger U.S. firms have been more able to participate in global markets than smaller firms because of their financial and managerial resources (Terpstra, Sarathy, and Russow 2006). The literature is not unanimous about the role of size in the success of firms; some researchers assert that large size helps, whereas others assert that it hurts.

There are several reasons larger firms might have greater success than smaller firms. First, larger firms have recourse to more resources or can commandeer more resources than smaller firms (Bonaccorsi 1992). For example, Coke was able to purchase the leading cola brand in India, Thums Up, to open its entry into India (Ramaswami and Namakumari 2004). Second, larger firms are more likely to possess a greater wealth of product-specific and marketing-specific knowledge than smaller firms. For example, Nestlé has a portfolio of 7695 brands to choose from and a huge organizational history of international expansion to help it exploit any new market that it enters (Parsons 1996). Third, larger firms are more capable of sustaining periods of negative performance on entry into a host country than smaller firms. Luo (1997) finds that size favors performance, even after controlling for mode of entry.

Conversely, the experience of many large firms shows that size is no guarantee for success. The recent withdrawal of Wal-Mart first from Korea and then from Germany is a case in point (*The Economist* 2006). Researchers have unearthed some explanations for this result. Large size diminishes organizational flexibility because of increasing bureaucracy (Hitt, Ireland, and Hoskisson 2003). This bureaucratic effect also impairs innovative ability (Chandy and Tellis 2000). In line with this finding, Cooper and Kleinschmidt (1985) show that export success is negatively correlated with firm size in the high-tech electronics industry. These arguments lead to our third research question:

> Q_3: Are smaller or larger firms more successful in entering merging markets?

Economic Distance

Economic distance is a measure of economic disparity between two countries. Firms find it easy to deal with host countries that are close in economic distance from their home country for several reasons. First, countries close in economic development have similar market segments that can afford to consume similar types of goods and services. Thus, knowledge about market demand transfers easily from home to host country. Second, countries close in economic development have similar physical infrastructure, such as airports, roadways, railways, and seaports. Thus, firms serving a host country with an infrastructure similar to the home country will enjoy efficiencies in its operations, thus lowering costs. Third, firms develop competencies or knowledge-based resources that are related to the markets they serve (Madhok 1997). These resources can be best leveraged in countries that are similar in economic development because the skills learned in one market can be replicated in or adapted to the new markets. Firms entering countries that are widely different economically from their home country need to adjust to the new market conditions, thus reducing their likelihood of success (Dunning 1998). These arguments suggest our fourth research question:

> Q_4: Does entry success decrease with greater economic distance?

Cultural Distance

Consumers are not driven by economic considerations alone. The underlying cultural dimensions of a society affect its consumption pattern beyond what economic laws predict (De Mooij 2004). "Culture" is usually defined as shared values and meanings of the members of a society. It affects not only the underlying behavior of customers in a market but also the execution and implementation of marketing and management strategies (Kogut and Singh 1988). For example, cultural distance affects how well partners in a joint venture interact over the cultural divide. Thus, cultural distance has a direct impact on the effectiveness of the entry.

Evidence of failures caused by insensitivity to cultural differences abounds. The much-discussed troubles of Euro Disney provide a classic example of how Disney executives failed to adjust for the cultural differences between the United States and Europe. Cultural differences affect several aspects of consumer behavior as well as a firm's marketing mix. It affects not only the attribute levels of products (Leclerc, Schmitt, and Dube 1994) and the efficiency of the marketing programs (Tse, Vetinsky, and Wehrung 1988) but also how customers derive meanings about the brand or product. Mistakes arising from misunderstandings of brand names are legion.

The tendency of firms to start their international marketing activities in countries similar to their own is another example of how culture influences market entry. Several studies have shown that the sequential path of internationalization is determined by cultural distance to enhance the chances of successful entry (Czinkota 1982). Firms usually begin internationalizing by entering countries that are culturally close to them. For example, Toyota began exporting by first selling to the Southeast Asian countries (Terpstra, Sarathy, and Russow 2006). In addition to geographic proximity, cultural similarities may also lead U.S. firms to trade with Canada, European countries to trade with one another, and Japanese firms to focus on Asia (Johansson 2006). Frankel and Rose (2002) show that linguistic similarity is a far more powerful determinant of the volume of trade between countries than economic factors, such as a common currency. Barkema, Bell, and Pennings (1996) also show that cultural barriers "punctuate" organizational learning, lowering firms' longevity in countries with greater cultural distance. These arguments suggest our fifth research question:

> Q_5: Does success into emerging markets decrease with greater cultural distance?

Country Risk

Erb, Harvey, and Viskanta (1995) define "country risk" as uncertainty about the environment, which has three sources: political, financial, and economic. Political risk is the risk that laws and regulations in the host country will be changed adversely against a foreign firm. These could be of a regulatory nature, such as the imposition of tariffs, or of a political nature, such as unrest caused by pressure groups (Spar 1997). At its severest, political risks may cause confiscation of assets without adequate compensation (Hawkins, Mintz, and Provissiero 1976).

Financial and economic risks manifest in several ways. They could take the form of (1) recessions or market downturns, (2) currency crises, or (3) sudden bursts of inflation. Most of these factors arise from imbalances in the underlying economic fundamentals of the host country, such as a balance of payment crisis. Recessions result from business cycles inherent in any economy (Lucas 1987). The origins of currency crises could be a progressively deteriorating trade imbalance (e.g., India in the late 1980s) or a loss of faith by the international financial system on the country's ability to meet its international debt obligations (e.g., Argentina in 2001). Whatever the source of the problem, a fall in the currency rate will lead to a fall in revenues and profits (Shapiro 1985). Differential inflationary pressures between the home and the host country could also pose a risk. Inflation directly affects the price–demand structure of a firm. It can also affect the firm indirectly through its adverse affects on exchange rates (Erb, Harvey, and Viskanta 1995; Frankel and Mussa 1980).

Country risk can reduce entry success in emerging markets in two ways. First, it can cause firms to lose money suddenly, precipitating a financial crisis. Consider P&G in Russia. Its "optimistic projections of Russia were shattered on a single day in the summer of 1998" (Dyer, Dalzell, and Olegario 2004, p. 336). The sudden devaluation of the ruble on August 17, 1998, triggered a deep financial crisis as the annual projected dollar revenues shrank to half, far below P&G's ability to service debts. A more serious problem was the uncertainty over how long the crisis would last. Second, high country risk and past experiences of risk can lead firms to underinvest or delay investments, resulting in lower success over time. Unilever was cautious and delayed entry into China, "especially in view of the past difficult experiences with the Soviet Union" (Jones 2005, p. 160), another high-risk country. These arguments suggest our sixth research question:

> Q_6: Does success of entry into emerging markets decrease with country risk?

Country Openness

The term "openness" refers to the lack of regulatory and other obstacles to entry of foreign firms. Openness could either increase or decrease entry success. On the one hand, openness could increase success for three reasons. First, it stimulates demand by increasing the variety of products offered for sale in the market. Second, it increases competition on quality and thus improves the level of quality supplied. Third, as the economy opens up, competition increases efficiency and lowers prices, resulting in further increases in demand. Consider the Indian automotive industry. Until the early 1980s, the protected local market was dominated by two highly inefficient players—Hindustan Motors (HM) and Premier Auto Limited (PAL)—which offered just 2 basic car models, priced at approximately $20,000. The government allowed Suzuki to set up a joint venture in 1983. This increased the number of car models in the Indian market to 3, and the quality of all cars on the market, including those from HM and PAL, improved dramatically. In 1992, the remaining barriers for foreign firms were lifted. Since then, 30 car models have been sold in India. Prices in all segments have steadily declined by 8%–10% a year, and the industry has tripled in size. The liberalization of the Indian telecommunications industry and the resultant boom in the sales of cell phones are other examples of how openness spurred growth in demand (Ramaswamy and Namakumari 2004). Evidence from China also shows that "growth acceleration has been associated with the opening of markets" (Naughton 2007, p. 7).

On the other hand, an open economy is a double-edged sword. Although openness makes entry easier for a target firm, it increases competition from other new foreign entrants. Increasing competition affects market success in several ways. First, even a small degree of competition is enough to lower prices significantly (Wallace 1998). Thus, competition keeps margins low, permitting only the most efficient to survive. Second, competition increases costs of purchases, the hiring of talent, and the marketing of products and services. Competitive pressures are a reason firm profitability has been shown to be lower for international markets than for domestic markets (Gestrin, Knight, and Rugman 2001). Third, competition causes firms to lose leadership if they make any strategic mistakes, such as targeting the wrong segment or pricing the product too high, both of which are common mistakes in entering emerging markets. Competitors are quick to pounce on any mistake and prevent firms from recovering lost ground. Thus, increasing openness increases competition and decreases success. These arguments suggest our seventh research question:

> Q_7: Does success of entry into emerging markets increase or decrease with country openness.

Summary

The prior sections show how three firm-level variables (mode of entry, timing of entry, and size) and four country-level variables (economic distance, cultural distance, risk, and openness) can affect the success or failure of a firm that is entering an emerging market. Next, we try to answer these questions through a historical analysis of entry into China and India.

Empirical Evidence

We carry out a historical analysis of market entry in two of the largest emerging markets to answer the research questions. We consider only the entry of firms that were not already set up in the years immediately before 1978 for China and 1991 for India. Historical analysis involves carefully assembling, critically examining, and summarizing the records of the past (Golder and Tellis 1993). This method is well suited for our purpose because it is based on neutral observers and factual data recorded at the time the success or failure of a firm's entry occurs. Historical analysis provides a powerful means of understanding marketing phenomena by recreating markets as they evolved (Golder 2000). It also responds to the call for historical research in this area (e.g., Jones and Khanna 2006). In particular, Mitra and Golder (2002, p. 382) recommend "longitudinal, archival-based studies of relative success of companies in multiple markets." This section presents the measures, procedure, sampling, and model of the empirics.

Measures

This subsection discusses the measures for the dependent variable and the seven independent variables: entry mode, entry timing, firm size, economic distance, cultural distance, and openness.

Dependent variable: success (or failure). Perhaps the most contentious issue in studying success and failure of international market entry is to define and measure it. This is so because firms do not divulge the internal parameters and measurements of success. Attempts to ascertain this by the survey method lead to the well-known self-reporting bias (Golder and Tellis 1993). In addition, success is a time-dependent phenomena, and at any given time, it may only be partial (Luo 1998). To circumvent this problem, researchers have used multiple measures of success, such as market share and profitability (Pan, Li, and Tse 1995), hazard rates (Li 1995), and timing (Luo 1998).

To arrive at an objective and comprehensive measure that can discriminate degrees of success, we used a content analysis of articles from several sources reporting on the performance of firms entering into China and India, and we derived numerical ratings. For the content analysis, we first developed a set of terms that reviewers use to describe success or failure of market entry. We then grouped these terms into five levels of increasing success, assessed on a five-point scale (see Appendix A). This graded measure of success enables us to measure degrees of success.

Entry mode. Anderson and Gatignon (1986) show how entry strategies can be categorized on the basis of the degree to which they allow a firm to control its entry into foreign markets. They categorize entry strategies as possessing low, medium, and high control over the firm's strategy. To calibrate the varying degrees of control, we used a six-point ordinal scale ranging from 1 ("low-control entry mode") to 5 ("high-control entry mode"), as follows: exports (1), alliances (2), franchise (3), joint ventures (4), equity joint ventures (4.5), and wholly owned subsidiaries (5).

Mixed entry modes, such as contract manufacturing, can be understood as a hybrid of existing modes. Idiosyncratic variations of the traditional entry modes, such as wet or dry licenses (see Luo 2000, p. 284), can also be defined within the scope of our scale. Firms with two entry modes for different products are considered two separate entries.

Timing. Our measure of timing is the number of years between a firm's market entry and the year of first deregulation by the host country. For China, we took 1978 as the first year of deregulation, and for India, we took 1991.

Firm size. To measure size of the firm, we used the year-end sales of the firm in the year of entry into the host country.

Economic distance. To measure economic distance, we followed the work of Mitra and Golder (2002). Thus:

$$(1)\qquad ED_{smt} = |GNP_{st} - GNP_{mt}| + |G\hat{N}P_{st} - G\hat{N}P_{mt}| + |Infra_{st} - Infra_{mt}| + |Popdensity_{st} - Popdensity_{mt}|,$$

where ED_{smt} is the economic distance between the host country s and the home country m in year t; $GNP_{st,\,mt}$ and $G\hat{N}P_{st,\,mt}$ are the log of aggregate and per capita gross national product (GNP) for host country s and home country m, respectively, in year t; $Infra_{st,\,mt}$ are the kilometers of road per square kilometer for host country s and home country m, respectively, in year t; and $Popdensity_{st,\,mt}$ are the population densities for host country s and home country m, respectively, in year t.

To capture the size of demand for a firm's goods in a host country, we used per capita GNP (Loree and Guisinger 1995). However, although per capita GNP provides a suitable measure for consumer goods, it does not give us a good measure for industrial products. To correct for this limitation, we used the aggregate GNP of the host country (Terpstra, Sarathy, and Russow 2006). We measured these variables in the year of entry and converted to their equivalent dollar values on the basis of year-end dollar exchange rates.

Cultural distance. We employed the measure of cultural distance between the host and the home countries from Hofstede's (1991) four cultural dimensions: power distance, individualism–collectivism, masculinity–femininity, and uncertainty avoidance. Following Kogut and Singh (1988), we collapsed the individual scores into a single number by taking the Euclidian distance of the four dimensions as follows:

$$CD_{smt} = \sum_{j=1}^{4} \sqrt{(D_{jst} - D_{jmt})^2}, \tag{2}$$

where CD_{smt} is the country distance score between host country s and home country m in year t, D_{jst} is the score on dimension j for host country s, and D_{jmt} is the score on dimension j for home country m both measured in year t. This measure of cultural distance has a long history of use in both the international marketing and the strategy literature (Mitra and Golder 2002).

Country risk. Our measure of country risk needs to capture political, regulatory, and economic sources of risk (Erb, Harvey, and Viskanta 1996; Simon 1984). Although several commercial agencies measure each of these components of country risk using proprietary methods, researchers in finance have shown that the ones used by the *International Country Risk Guide* possess the greatest forecast accuracy (Erb, Harvey, and Viskanta 1997). This measure of country risk is based on a multidimensional measure for each component of country risk through political, financial, and economic risk (for details, see Appendix A). We reverse-coded country risk relative to the United States, which has the highest score and the lowest risk.

Openness. Our measure of openness is based on the fraction of foreign direct investment as a function of the host country's GDP. We compiled this measure from the annual statistical surveys of China and India.

Procedure

The data for this study are a unique compilation from several sources (see Appendix B). The primary source for information about market entry and market success is from electronic sources, such as LexisNexis and ABI/INFORM. Golder and Tellis (1993) show that archival data must meet the following criteria to ensure validity:

- •*Competence*: the capability of the informant to report correctly,
- •*Neutrality*: the lack of vested interest by the informant of the report,
- •*Reliability*: a long record for undisputed good reporting by the informant,
- •*Corroboration*: confirmatory evidence from a similar source, and
- •*Contemporaneity*: proximity of the time of the report to that of the event.

The competence criterion is met because the reports are by well-known sources and are from the time frame when the firms entered the host country. The objectivity criterion is satisfied because neutral commentators wrote the stories. The reliability criterion is satisfied because the sources are all reputable. The corroboration criterion is satisfied because at least two data sources are used to complete the details for each firm. Contemporaneity is satisfied because the electronic search engines sorted the articles with the oldest first to ensure that the reports closest to the event are included in the sample. We collected additional articles when necessary so that the data on success and failure would meet these criteria. We used hard-copy sources, such as books and country reports (e.g., International Monetary Fund country reports), to supplement the electronic sources. The period of the data coincides with the period in Study 1. The following is a step-by-step elaboration of this technique:

1. Locate articles on entry into China and India using key words.
2. Extract and save articles from Step 1, or when applicable, obtain hard copies. Extract information on firm names and enter this into a spreadsheet.
3. Extract phrases about the success or failure of the entry and record them in the same spreadsheet as that in Step 2.
4. Compile additional information on the mode of entry, performance of the firm in the host country, and year of entry by focusing the search on the firm and expanding the key words.

We study the information collected to arrive at the five-point scale for success and failure (see Appendix C). We recruited and trained two MBA students as research assistants for the study. The research assistants evaluated the language of each review using the five-point scale. They then converted the review into a numerical rating of success. We instructed them to treat the scale as continuous from 1 to 5. The assistants were allowed to consult the authors for any interpretive difficulties. We used the average rating from the two assistants for the analysis. The correlation coefficient of the coding between the two research assistants was .78. The interrater reliability, as measured by Cronbach's alpha, was .88. The research assistants were within one count of each other for 88% of the cases. All these statistics compare favorably with those of Chandy and colleagues (2001).

We retrieved and coded data on entry mode from the archival data. Data on sales at the time of entry were collected and recorded in millions of local currency primarily from three sources: COMPUSTAT tapes from the Wharton Research Data Services for U.S. firms and from firms' Web sites and Mergent Online database for non-U.S. firms. We converted all sales data into U.S. dollars for analysis. We collected data on cultural dimensions from the work of Hofstede (1991, 2001). We obtained economic measures from the *International Financial Statistics Yearbook*, a compilation of annual national statistics prepared by the International Monetary Fund. This was also the source of foreign-exchange rates needed to convert sales figures and GNP data denominated in local currency into U.S. dollars. We used the year-end market exchange rates whenever available and government-nominated rates elsewhere. We acquired data on country risk for each year of interest from the *International Country Risk Guide*, which is available online from the Web site maintained by the PRS Group (http://www.prsgroup.com).

Sample

Beginning with 192 entries of firms into China and India that met with the outlined criteria, we found that 128 entries were into China and 64 were into India. The number of entries we found for India is substantially less given that India's major economic reforms took place 13 years after China's reforms began. In 9 cases, we could not obtain information on the mode of entry of the firms, and in another 11 cases, the exact nature of success or failure was not clear. For 9 additional cases, sales data were missing. These are non-U.S. firms that entered in the 1980s and early 1990s for which we could not obtain any records in the public domain. Missing sales values are replaced with the mean dollar sales value of the entire sample. Thus, the usable sample is 168 cases.

Model

To answer the research questions, we estimated the following regression model:

$$
\begin{aligned}
(3)\quad \text{Success}_{ismt} = {} & \beta_1 \times \text{Entry mode}_{ism} + \beta_2 \times \text{Timing}_{ismt} + \beta_3 \\
& \times \text{Size}_{it} + \beta_4 \times \text{Cultural distance}_{smt} + \beta_5 \\
& \times \text{Economic distance}_{smt} + \beta_6 \\
& \times \text{Country risk}_{st} + \beta_7 \times \text{Openness}_{st} + \beta_8 \\
& \times \text{India} + \beta_9 \times \text{Entry mode}_{ism} \times \text{India} + \beta_{10} \\
& \times \text{Timing}_{ismt} \times \text{India} + \beta_{11} \times \text{Size}_{it} \\
& \times \text{India} + \beta_{12} \times \text{Economic distance}_{smt} \\
& \times \text{India} + \beta_{13} \times \text{Cultural distance}_{smt} \\
& \times \text{India} + \beta_{14} \times \text{Country risk}_{st} \times \text{India} + \beta_{15} \\
& \times \text{Openness}_{st} \times \text{India} + \varepsilon_{ismt},
\end{aligned}
$$

where i is a subscript for firm, s is a subscript for host country, m is a subscript for home country, and t is a subscript for time. Success is the success rating from 1 to 5, Entry mode is the categorical variable specifying the mode of entry chosen by the firm, Timing is the number of years between the year of a firm's entry and the start of economic reforms in the host country, Size is the logarithm of dollar value of sales (in million) in year of entry, Economic distance is given by Equation 1, Cultural distance is the difference between the host and the home countries in the composite measure we calculated from Hofstede's (1991) individual dimensions (Equation 2), Country risk is the overall country risk of the host country, Openness is a measure of the degree of participation of foreign firms in the host country, India is a dummy variable, β_1–β_{15} are coefficients to be estimated, and ε_{ismt} is an error term initially assumed to be i.i.d. normal. To ascertain the heterogeneity of coefficients over China and India, we include a dummy variable for and interaction terms of India with each of the key independent variables.

Results

Descriptive Statistics

Table 1 provides the descriptive statistics for our sample of firms. The table shows that the dominant mode of new entry into China (1978–2005) and India (1991–2005) is the joint venture (41%), followed by the wholly owned subsidiary (33%) and the equity joint venture (10%). Exports, licensing, and franchising make up 4%, 7%, and 5%, respectively. Of the entering firms, 56% were from North America (United States and Canada), 23% were from Europe, and 21% were from Southeast Asia (including Australia and New Zealand).

Model Estimates

Table 2 reports results of the estimation of the model in Equation 3. To ascertain the effect of multicollinearity, if any, we provide estimates of running a simple regression with each independent variable alone (Columns 3 and 4), a full model with all important interaction terms (Columns 5 and 6), and a reduced model after dropping insignificant terms (Columns 7 and 8). Note that all the main effects are significantly different from 0. In addition, the main effect of India and three of the interaction terms with India (timing, economic distance, and cultural distance) are significantly different from 0, suggesting that these three drivers hold differently for India and China. Conversely, four of the interaction effects with India (entry mode, size, risk, and openness) are not significantly different from 0, suggesting that these four drivers hold equally well for China and India. The main effect for India is negative and significant, suggesting

TABLE 1
Descriptive Statistics

Variable	Number	Percentage
1. Mode of Entry		
a. Exports and branch subsidiaries	7	4
b. Licenses	12	7
c. Franchises and agreements	10	5
d. Joint ventures	75	41
e. Equity joint ventures	18	10
f. Wholly owned subsidiaries	61	33
2. Country of Origin of Firms		
a. North America	108	56
b. Europe	43	23
c. Southeast Asia, Australia, and New Zealand	41	21
3. Type of Industry		
a. Consumer nondurable	54	28
b. Consumer durable	86	46
c. Service	29	15
d. Industrial	23	11

TABLE 2
Regression of Success on Its Determinants

		Simple Regression Estimates		Full-Model Estimates		Reduced-Model Estimates	
Dependent Variable (1)	Question (2)	Estimate (3)	t (4)	Estimate (5)	t (6)	Estimate (7)	t (8)
Intercept				23.981	2.767**	20.040	3.031**
Entry mode	Q_1	.261	2.946**	.273	2.396**	.288	3.521**
Timing	Q_2	–.043	–2.662**	.236	1.531	.212	1.620
Size	Q_3	–.121	–2.041**	–.141	–1.965*	–.146	–2.740**
Economic distance	Q_4	–.002	–1.991**	–.008	–2.725**	–.007	–3.063**
Cultural distance	Q_5	–.017	–2.168**	.008	.696	.006	.634
Country risk	Q_6	.042	3.003**	.029	.848	.048	2.029**
Openness	Q_7	–.038	–2.134**	.155	.164	–.085	–2.27**
India		–.630	–2.752**	–34.765	–3.165**	–32.295	–3.489**
Entry mode × India				.123	.680		
Timing × India				–.351	–1.999*	–.334	–2.193**
Size × India				.002	.016		
Economic distance × India				.011	2.620**	.011	3.110
Cultural distance × India				–.076	–2.457**	–.076	–2.670**
Country risk × India				.042	.816		
Openness × India				–.256	–.271		
Adjusted R^2				29.02%		28.43%	
F				4.279***		5.814***	
N				168		168	

*$p < .1$.
**$p < .05$.
***$p < .001$.

that, in general, entry into India has been less successful than entry into China. Moreover, this effect is robust, holding equally strongly across all three specifications. The R-square is approximately 29%, which compares well with other studies (Gatignon and Anderson 1988; Pan, Li, and Tse 1999).

Drivers of Success or Failure

With reference to Q_1, the positive and highly significant coefficient for entry modes shows that those that allow for higher control tend to be more successful. The effect is robust and holds for all three model specifications (Columns 3–8 in Table 2) and holds equally strongly for China and India.

With reference to Q_2, the negative and highly significant coefficient for entry timing (Columns 3 and 4) shows that firms that entered earlier were more successful. However, this effect seems to hold only for India, as indicated by the negative and significant interaction term for India (Columns 5 and 6).

With reference to Q_3, the negative and significant coefficient for size shows that smaller firms have greater success in emerging markets. The effect is robust; it holds across all three specifications (Columns 3–8) and is equally strong for China and India (Columns 5 and 6).

With reference to Q_4, the negative and significant coefficient for economic distance shows that firms that enter host countries that are economically similar to the home country enjoy greater success. This effect is robust, and it holds across all three specifications (Columns 3–8). However, it is significantly weaker for India than for China (Columns 5 and 6).

With reference to Q_5, the negative and significant coefficient for cultural distance shows that firms that enter host countries that are culturally closer to the home country enjoy greater rates of success (Columns 3 and 4). However, this effect is not robust and does not hold in the presence of other variables. The effect holds in the expected direction only for India, as evidenced by the significant interaction effect with India (Columns 5 and 6).

With reference to Q_6, the positive and significant coefficient for country risk shows that greater risk of the host country leads to less success (Columns 3, 4, 7, and 8). Note that country risk is reverse coded, signifying higher scores for lower-risk countries. Moreover, this effect holds equally strong for China and India.

With reference to Q_7, the negative and significant coefficient for openness shows that greater openness lowers success (Columns 3, 4, 7, and 8). This effect holds equally strong for China and India.

Conclusion

Contribution

China and India are two of the largest emerging markets. They are growing quickly and are destined to rank among the top economies of the world in the next two or three decades. Firms are in a rush to enter these markets. However, the literature contains insufficient analysis of the drivers of success and failure of entry in these markets. Our study makes four distinct contributions in this area. First, we offer a richer definition of success and failure than prior studies. Second, we relate our measure of success to important causal drivers, which emerge from a vast body of interdisciplinary research over decades. The results show which of these drivers are most important and whether the importance is generalizable or pertains to only one country. Third, we focus on both the major emerging markets of China and India. Fourth, we use a research method that is rarely used in this domain—namely, historical analysis. The main conclusions from our study are the following.

- Success is greater for entry into China than for entry into India.
- Success is greater for smaller firms than for larger firms.
- Success is greater for entry into emerging markets with less openness and less risk and those that are economically close to the home market.
- Success is greater for firms that use a mode of entry with greater control.
- Joint ventures are the most popular mode of entry, accounting for 41% of entry modes.

Discussion

Perhaps the most surprising finding is that success is substantially and significantly lower in India than in China. One possible reason for this is the immense diversity of India, which is characterized by inconsistent policy across Indian states and pockets of varying demand across the India market. A second possible reason is that India had an early history of capitalism with many entrenched private firms and brand names. Thus, entrants had greater native competition in India than in China. A third reason could be that China's infrastructure has been substantially superior to India's, making operations much easier for new entrants.

A second surprising finding is that smaller firms tend to be more successful than larger firms in entering emerging markets. This result is contrary to research findings, which have shown that a larger firm size correlates with greater success (e.g., Anderson and Gatignon 1986; Luo 1997). An example may clarify this result. General Motors, the largest automaker in sales, and Toyota, the largest in market capital, have struggled in India, whereas smaller rivals, such as Hyundai, have been successful. An explanation for this result is that the mere size of resources itself may not be the chief factor behind success. Control of resources and how they are deployed may lie at the heart of success in China and India because these markets are characterized by rapid environmental changes that require continuous adaptability and learning (Yan 1998). Thus, small firms with a less bureaucratic burden may be able to adapt more quickly (Hitt, Ireland, and Hoskisson 2003). Indeed, researchers in international marketing have found that smaller firms, given their smaller budgets, tend to collect first-hand information rather than sponsor third-party data collection (Hollensen 2004). Another explanation is that larger firms may be more confident or even arrogant about their resources, strengths, and prior successes and therefore may not try as hard to succeed as smaller firms (Chandy and Tellis 2000).

A third surprising result is that the openness of markets reduces success in both China and India. Intuitively, openness suggests easier entry and, thus, easier success. However, what is often overlooked is that what is true for one entrant is also true for other entrants. Greater openness results in more firms from the same industries from multiple countries entering the fray. This competition puts downward pressure on margins, making it increasingly difficult for all firms to succeed. Thus, increasing openness increases competition and decreases success.

Consistent with this result, we find that earlier entrants enjoy greater success than later entrants, at least in India. This finding is consistent with prior studies (Pan and Chi 1999). Indeed, content analysis of archival reports of the reasons for success and failure shows that the speed of entry was mentioned 25 times in the reports. For example, P&G, which entered India much later than Unilever, does not have the market success of Unilever.

A strong finding of our study is that entry strategies that involve high levels of control (e.g., wholly owned subsidiaries) are more successful than those that involve low levels of control (e.g., licensing). For example, in China, FedEx, which operates as a wholly owned subsidiary, is more successful than UPS, which operates as a joint venture. Our results hold despite the possible entry restrictions on mode of entry that China and India have imposed. Restriction to entry usually forces firms to take low-control entry modes. However, we still have a large proportion of observations for high-control modes and find this variable to be highly significant.

Economic and cultural proximity between the home and the host country favors successful entry into emerging markets. For example, Charoen Pokphand Group, the Southeast Asian agribusiness conglomerate from Thailand, is more successful in China than Seagram, the agri-based firm of North America. The effect of cultural distance is far stronger in India than in China. Our content analysis of the archival reports indicates that one of the most frequently cited reasons (34 times) for success or failure in India is how well or poorly (respectively) the entrant adapts the product to the local culture. Surprisingly, even after several decades of international experience, many Western firms tend to impose Western consumption habits and production methods in emerging markets. For example, Kellogg initially failed to market cold breakfast cereal in India because of the strong Indian taste for hot breakfast foods.

Implications

This research has some important implications for entry into emerging markets. First, firms should consider not only the growth of emerging markets but also the success rates of prior entrants. In the case of the two countries under study, China seems to have a much higher success rate than India.

Second, the progressive opening of the economies of China and India does not mean that firms should wait to enter when entry gets easier. Easier entry applies to all firms, thus increasing competition. As China and India liberalize and deregulate even further, the increased competition will reduce success. Our data suggest that earlier entrants enjoy greater success. Thus, firms that enter later should be prepared for stiffer competition and probably less success.

Third, counter to widely held priors, small size itself should not deter firms from entering emerging markets. In contrast, large firms should not assume that past success and deep resources will necessarily guarantee success.

Fourth, firms should choose the entry mode that affords them the greatest degree of control when entering emerging markets. Doing so implies not taking on partners and alliances in the host country, which may add to the cost and difficulty of entry. However, the greater control provides the entrant with an opportunity to compete on its own unique strengths, monitor success and failure closely, and make changes in strategy as soon as necessary.

Fifth, when entering emerging markets, firms should consider targets that are close to their home country in terms of economic and cultural distance. In particular, firms from developing countries may be more successful in entering emerging markets than those from developed countries, if the emerging markets are close to them in cultural or economic distance. An example is the inroads made by Chinese and South Korean firms into the emerging markets of India and Brazil.

Limitations and Further Research

Our study has several limitations that could benefit from further research. First, analysis of disaggregate firm-level variables, such as the level of investment in manufacturing and marketing, could further enlighten the issues. Second, research on whether and which firms learn from their mistakes would be helpful. Third, more precise measures of culture are in order. The standard country-level measures, such as Hofstede's (1991, 2001) cultural distance, are at a too aggregate level and are static in nature and may not reflect the regional differences and temporal changes in large countries, such as China and India. Fourth, the evolution of a firm's fortunes over time could lead to greater insights into how the firm adjusts its strategies to exploit the opportunities presented by emerging markets. Fourth, whereas economic and cultural distance measures are proxies of firm knowledge, other drivers, such as experience in similar markets, may be important proxies of firm knowledge. Fifth, entering firms may have faced regulatory restrictions over their choice of entry mode, which may have restricted the full set of options normally available.

Appendix A
Details of Country Risk Calculation (from the *International Country Risk Guide*)

Political risk is calculated by assigning points to each of the following components: government stability, socioeconomic conditions, investment profile, internal conflict, external conflict, corruption, military in politics, religious tensions, law and order, ethnic tensions, democratic accountability, and bureaucracy quality.

Financial risk is calculated by assigning points to each of the following components: total foreign debt as a percentage of GDP, debt service as a percentage of exports of goods and services, current account as a percentage of exports of goods and services, international liquidity as months of import cover, and exchange rate stability as a percentage of change.

Economic risk is calculated by assigning points to each of the following components: real annual GDP growth, annual inflation rate, budget balance as a percentage of GDP, and current account as a percentage of GDP.

A composite country risk is produced by combining these three measures according to the following formula:

$$\text{CPFER} = .5\ (\text{PR} + \text{FR} + \text{ER}), \tag{A1}$$

where CPFER is a composite of political, financial, and economic risk ratings; PR is the total political risk indicators; FR is the total financial risk indicators; and ER is the total economic risk indicators. The highest overall rating (theoretically, 100) indicates the lowest risk, and the lowest rating (theoretically, 0) indicates the highest risk.

Appendix B
Sources

India

American Chamber of Commerce
Asiaweek
BBC
Business India Intelligence
BusinessWeek
The Economic Times (India)
The Economist
Harvard Business Review
India Brand Equity Foundation
India Today
McKinsey
The New York Times
The Telegraph
The Times of India
The Wall Street Journal
The Wall Street Journal (Asia)

China

AmCham News
AmCham News: China Briefs
American Chamber of Commerce
Asiaweek
Bain Consulting Company
BBC
Business Times
Business-China.com
China Business Insight
China Wire
China Bulletin
The China Business Review
The Economist
Fortune
Global News Wire
Harvard Business Review
McKinsey
Mintel's Global New Products Database
People's Daily
South China Morning Post
The Wall Street Journal (Asia)
Xinhua News Agency

Appendix C
Content Analysis Outline

The outline for evaluating success or failure of entry is given as follows:

Successful Entry: 5

- Making more margins than their global margin
- Market share leader
- Well-functioning partnership
- Above-average industry leadership
- Top three in industry profitability
- Top three in market share
- Exceeded investment criteria

Good Entry: 4

- Successfully selling
- Met investment criteria
- Increasing investments
- Growing shipments
- Rapidly evolved into a major force in the industry

Acceptable Entry: 3

- Hope to recover investment in time
- Entry awaiting removal of market restrictions
- Establish a beachhead
- Continuing operations

Poor Entry: 2

- No initial lead buyers
- Conflicting expectations
- Fail in system integration and optimization
- Struggled to make headway
- Underperformance
- Priced out
- Stiff competition
- Market restrictions
- Executives frustrated with entry

Failed Entry: 1

- Quit or withdrawal from market
- Break up with cessation of venture

REFERENCES

Anderson, Erin and Hubert Gatignon (1986), "Modes of Foreign Entry: A Transaction Cost Analysis and Propositions," *Journal of International Business Studies*, 17 (Fall), 1–25.

Arnold, David (2004), *The Mirage of Global Markets*. Upper Saddle River, NJ: Prentice Hall.

Barkema, Harry G., John H.J. Bell, and Johannes M. Pennings (1996), "Foreign Entry, Cultural Barriers, and Learning," *Strategic Management Journal*, 17 (2), 151–66.

Bonaccorsi, Andrea (1992), "On the Relationship Between Firm Size and Export Intensity," *Journal of International Business Studies*, 23 (4), 603–635.

Carpenter Gregory S. and Kent Nakamoto (1989), "Consumer Preference Formation and Pioneering Advantage," *Journal of Marketing Research*, 26 (August), 285–98.

Cateora, Philip R. and John L. Graham (2006), *International Marketing*, 13th ed. New York: McGraw-Hill.

Chandy, Rajesh K. and Gerard J. Tellis (2000), "The Incumbent's Curse? Incumbency, Size, and Radical Product Innovation," *Journal of Marketing*, 64 (July), 1–17.

———, ———, Deborah J. MacInnis, and Pattana Thaivanich (2001), "What to Say When: Advertising Appeals in Evolving Markets," *Journal of Marketing Research*, 38 (November), 399–414.

Cooper, Robert G. and Elko J. Kleinschmidt (1985), "The Impact of Export Strategy on Export Sales Performance," *Journal of International Business Studies*, 16 (Spring), 37–55.

Czinkota, Michael R. (1982), *Export Development Strategies: U.S. Promotion Policies*. New York: Praeger Publishers.

Dasgupta, Jayatri and Sanjib Dutta (2004), "Unilever's Strategies in China," ICFAI Center for Management Research Case No. BSTR 131.

De Mooij, Marieke (2004), *Consumer Behavior and Culture: Consequences for Global Marketing and Advertising*. Thousand Oaks, CA: Sage Publications.

Dunning, John H. (1988), "The Eclectic Paradigm of International Production: A Restatement of Some Possible Extensions," *Journal of International Business Studies*, 19 (Spring), 1–31.

——— (1998), "Location and the Multinational Enterprise: A Neglected Factor?" *Journal of International Business Studies*, 29 (1), 45–66.

Dyer, Davis, Fredrick Dalzell, and Rowena Olegario (2004), *Rising Tide*. Boston: Harvard Business School.

The Economist (2006), "Heading for the Exit," (August 5), 54.

Erb, Claude, Campbell R. Harvey, and Tada Viskanta (1995), "Inflation and World Equity Selection," *Financial Analysts Journal*, 51 (November–December), 28–42.

———, ———, and ——— (1996), "Political Risk, Financial Risk and Economic Risk," *Financial Analysts Journal*, 52 (6), 28–46.

———, ———, and ——— (1997), *Country Risk in Global Financial Management*. Charlottesville, VA: Association for Investment Management and Research.

Fishman, Ted (2005), *China*Inc.: How The Rise of the Next Superpower Challenges America and the World*. New York: Simon and Schuster.

Frankel, Jacob A. and Michael L. Mussa (1980), "The Efficiency of Foreign Exchange Markets and Measures of Turbulence," *American Economic Review*, 70 (2), 374–81.

——— and Andrew Rose (2002), "An Estimate of the Effect of Common Currencies on Trade and Income," *Quarterly Journal of Economics*, 117 (2), 437–66.

Fujikawa, Yoshinori and John A. Quelch (1998), "Sony Corporation: Car Navigation Systems," Harvard Business School Case No. 9-9597-032.

Gatignon, Hubert and Erin Anderson (1988), "The Multinational Corporation's Degree of Control Over Foreign Subsidiaries: An Empirical Test of a TransactionCost Explanation," *Journal of Law, Economics, and Organization*, 4 (2), 89–120.

Gestrin, Michael V., Rory F. Knight, and Alan M. Rugman (2001), *Templeton Performance Index*. Oxford: University of Oxford.

Gielens, Katrijn and Marnik G. Dekimpe (2007), "The Entry Strategy of Retail Firms into Transition Economies," *Journal of Marketing*, 71 (April), 196–212.

Gillespie, Kate, Jean-Pierre Jeannet, and H. David Hennessy (2007), *Global Marketing*, 2d ed. Boston: Houghton Mifflin.

Golder, Peter N. (2000), "Historical Method in Marketing Research with New Evidence on Long-Term Market Share Stability," *Journal of Marketing Research*, 37 (May), 156–72.

——— and Gerard J. Tellis (1993), "Pioneer Advantage: Marketing Logic or Marketing Legend?" *Journal of Marketing Research*, 30 (May), 158–70.

Hawkins, Robert G., Norman Mintz, and Michael Provissiero (1976), "Government Takeovers of U.S. Foreign Affiliates," *Journal of International Business Studies*, 7 (1), 3–16.

Hawksworth, John (2006), *The World in 2050: How Big Will the Major Emerging Market Economies Get and How Can the OECD Compete?* New York: PricewaterhouseCoopers.

Hitt, Michael A., R. Duane Ireland, and Robert E. Hoskisson (2003), *Strategic Management: Competitiveness and Globalization*, 6th ed. Mason, OH: South-Western.

Hofstede, Geert (1991), *Culture and Organizations: Software of the Mind*. New York: McGraw-Hill.

——— (2001), *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, 2d ed. Thousand Oaks, CA: Sage Publications.

Hollensen, Svend (2004), *Global Marketing: A Decision-Oriented Approach*, 3d ed. Upper Saddle River, NJ: Prentice Hall.

India Brand Equity Foundation (2005), "Fortune 500 Companies in India: Success Stories," report, (July 7).

Isobe, Takehiko, Shige Makino, and David B. Montogmery (2000), "Resource Commitment, Entry Timing, and Market Performance of Foreign Direct Investments in Emerging Economies: The Case of Japanese International Joint Ventures in China," *Academy of Management Journal*, 43 (3), 468–84.

Johansson, Johnny K. (2006), *Global Marketing*, 4th ed. Homewood, IL: McGraw-Hill/Richard D. Irwin.

Jones, G. (2005), *Renewing Unilever*. Oxford: Oxford University Press.

——— and Tarun Khanna (2006), "Bringing History (Back) into International Business," *Journal of International Business Studies*, 37 (July), 453–68.

Kogut, Bruce and Harbir Singh (1988), "The Effect of National Culture on the Choice of Entry Mode," *Journal of International Business Studies*, 19 (3), 411–32.

Krugman, P.R. (1980), "Scale Economies, Product Differentiation, and the Pattern of Trade," *American Economic Review*, 70 (5), 950–59.

Leclerc, France, Bernd H. Schmitt, and Laurette Dube (1994), "Foreign Branding and Its Effects on Product Perceptions and Attitudes," *Journal of Marketing Research*, 31 (May), 263–70.

Li, Jiatao (1995), "Foreign Entry and Survival: Effects of Strategic Choices on Performance in International Markets," *Strategic Management Journal*, 16 (5), 333–51.

Loree, David W. and Stephen E. Guisinger (1995), "Policy and Non-Policy Determinants of U.S. Equity Foreign Direct Investment," *Journal of International Business Studies*, 26 (2), 281–99.

Lucas, Robert E. (1987), *Models of Business Cycles*. New York: Basil Blackwell.

Luo, Y. (1997), "Performance Implications of International Strategy: An Empirical Study of Foreign-Invested Enterprises in China," *Group and Organization Management*, 22 (1), 87–116

——— (1998), "Timing of Investment and International Expansion Performance in China," *Journal of International Business Studies*, 29 (2), 391–407.

——— (2000), *Multinational Corporations in China*. Hendon, VA: Copenhagen Business School Press.

——— (2001), "Determinants of Entry in an Emerging Economy: A Multilevel Approach," *Journal of Management Studies*, 38 (3), 443–72.

Madhok, A. (1997), "Cost, Value and Foreign Market Entry Mode: The Transaction and the Firm," *Strategic Management Journal*, 18 (1), 39–62.

Mitchell, W. (1999), "Dual Clocks: Entry Order Influences on Incumbent and Newcomer Market Share and Survival When Specialized Assets Retain Their Value," *Strategic Management Journal*, 12 (2), 85–100.

Mitra, Debanjan and Peter N. Golder (2002), "Whose Culture Matters? Near-Market Knowledge and Its Impact on Foreign Market Entry Timing," *Journal of Marketing Research*, 39 (August), 350–65.

Naughton, Barry (2007), *The Chinese Economy: Transitions and Growth*. Cambridge, MA: MIT Press.

Pan, Yigang and Peter S.K. Chi (1999), "Financial Performance and Survival of Multinational Corporations in China," *Strategic Management Journal*, 20 (4), 359–74.

———, Shaomin Li, and David K. Tse (1999), "The Impact of Order and Mode of Market Entry on Profitability and Market Share," *Journal of International Business Studies*, 30 (1), 81–104.

Parsons, Andrew J. (1996), "Nestle: The Visions of Local Managers," *The McKinsey Quarterly*, 2 (2), 5–29.

Porter, Michael E. (1990), *Competitive Advantage of Nations*. New York: The Free Press.

Ramaswamy, V.S. and S. Namakumari (2004), *Marketing Management: Global Perspective: The Indian Context*, 3d ed. New Delhi, India: Macmillan.

Root, Franklin R. (1994), *Entry Strategies for International Markets*. New York: Lexington Books.

——— and A. Ahmed (1979), "Empirical Determinants of Manufacturing Direct Foreign Investment in Developing Countries," *Economic Development and Cultural Change*, 27 (July), 751–54.

Shapiro, Alan C. (1985), "Currency Risk and Country Risk in International Banking," *Journal of Finance*, 3 (July), 881–91.

Simon, Jeffrey D. (1984), "A Theoretical Perspective on Political Risk," *Journal of International Business Studies*, 15 (3), 123–43.

Spar, Debora (1997), "Note on Political Risk Analysis," Harvard Business School Note No. 9-798-022.

Tellis, Gerard J. and Peter N. Golder (2001), *Will & Vision: How Latecomers Grow to Dominate Markets*. New York: McGraw-Hill.

Terpstra, Vern, Ravi Sarathy, and Llyod Russow (2006), *International Marketing*. Garfield Heights, OH: North Coast Publishers.

Tse, David K., Kam-hon Lee, Ilan Vertinsky, and Donald A. Wehrung (1988), "Does Culture Matter? A Cross-Cultural Study of Executives' Choice, Decisiveness, and Risk Adjustment in International Marketing," *Journal of Marketing*, 52 (October), 81–95.

Tunistra, Fons (2000), "Why America's P&G Is Way Ahead of Europe's Unilever in China," *Asiaweek*, (January 21), 26.

Vernon, Raymond, Louis T. Wells, and Subramanian Rangan (1996), *The Manager in the International Economy*. Upper Saddle River, NJ: Prentice Hall.

Vietor, Richard H.K. (2007), *How Countries Compete*. Boston: Harvard Business School Press.

Wallace, William McDonald (1998), *Post Modern Management: The Emerging Partnership Between Employees and Stockholders*. Westport, CT: Quorom Books.

Wilson, Dominic and Roopa Purushothaman (2003), "Dreaming with BRICs: The Path to 2050," Global Economics Paper No. 99.

Yan, Rick (1998), "Short-Term Results: The Litmus Test for Success in China," *Harvard Business Review*, 76 (September–October), 61–75.

Zhao, Hongxin, Yadong Luo, and Taewon Suh (2004), "Transaction Cost Determinants and Ownership-Based Entry Mode Choice: A Meta-Analytical Review," *Journal of International Business Studies*, 35 (6), 524–44.

Koen Pauwels & Allen Weiss

Moving from Free to Fee: How Online Firms Market to Change Their Business Model Successfully

Moving from free to "free and fee" for any product or service represents a challenge to managers, especially when consumers have plenty of free alternatives. For one online content provider, this article examines (1) the sources of long-term revenue loss (through attracting fewer free subscribers) and (2) how the firm's marketing actions affect its revenue gains (through attracting paid subscribers). The authors quantify revenue loss from several sources, including the direct effects of charging for part of the online content and the reduced effectiveness of search-engine referrals and e-mails. The analysis suggests several managerial implications. Managers should focus their price promotions on stimulating new monthly subscriptions, rather than the current promotional focus on stimulating new yearly contracts. In contrast, e-mail and search-engine referrals appear to be effective at generating yearly subscriptions. Meanwhile, free-to-fee conversion e-mail blasts are a double-edged sword; they increase subscription revenue at the expense of advertising revenue. Finally, further analysis shows that the move was preceded by the buildup of momentum in new free subscriptions, which appears to be beneficial for the move's success. The decomposition and comparison of the sources of revenue loss versus gains reveals several trade-offs facing companies moving from free to free and fee.

Keywords: free to fee, online content, price, marketing communication, search-engine referrals, time-series analysis, vector-autoregressive models

A decade ago, Peterson, Balasubramanian, and Bronnenberg (1997) heralded the Internet as a promising new channel for products and services that are low outlay, are frequently purchased, and have a value proposition of an intangible or informational nature. Such content providers have indeed become popular with consumers (Saba 2005). Recent research in marketing has made significant advances in the understanding of various Internet marketing issues, including Web-browsing behavior (Bucklin and Sismeiro 2003), search-engine visits (Telang, Boatwright, and Mukhopadhyay 2004), and recommendation agents (Cooke et al. 2002; Diehl, Kornish, and Lynch 2003). However, a vexing problem for practitioners has remained virtually untouched by formal analysis—namely, how firm performance is affected by moving from free to free and fee (i.e., from offering all content for free to charging for at least some of it).[1] A prominent fear among content providers is the likely loss of customers to the many still-free competitors. Moreover, the impact of marketing actions to stimulate paid subscriptions is uncertain given the general consensus among Web surfers that "content is free" (Dyson 1995). Empirical findings on these issues are scarce because firms are uneasy sharing data with academic researchers for confidentiality and competitive reasons.

Consumers like getting products and services for free. Beyond the absence of monetary costs, a free good is free of trade-offs and reduces consumer psychological costs (Ariely and Shampan'er 2004). Thus, being free offers an incentive to try a new product. However, it may pose a problem if the company decides that customers should begin paying. The issue of moving from free to fee matters in many industries, including entertainment content providers and the publishing industry, which represented $240 billion in 2004, according to the U.S. Census Bureau (see http://www.census.gov/svsd/www/services/sas/sas_data/51/2006_NAICS51.pdf). After the Internet "bubble" burst—this period was characterized by free online content that was supported by investor funding and high advertising income—many companies tried charging for their Web content, with varying levels of success.[2] Research by economists on the pricing of information goods (e.g., Bakos and Brynjolfsson 2000; Jain and Kannan 2002; Varian 1995) identifies various pricing models that can be employed to

---

[1] Throughout this article, we use "the move from free to fee" as a shorter term for "the move from free to free and fee."

[2] Insight into this issue appears to be crucial for content providers, whose increasing investment in producing high-quality content has put a strain on an advertising-only business model. At the same time, this traditional advertising income faces new competition in several forms, such as search-engine advertisements and classified advertisements from successful entrants (e.g., Google Base, Craigslist, Wikipedia). Therefore, getting subscribers to pay for at least some valued content may represent a more stable and sustainable strategy in the long run.

Koen Pauwels is Associate Professor of Business Administration, Tuck School of Business, Dartmouth College (e-mail: koen.h.pauwels@dartmouth.edu). Allen Weiss is Professor of Marketing, Marshall School of Business, University of Southern California (e-mail: amweiss@marshall.usc.edu).

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (May 2008), 14–31

guide these decisions. These include usage-based pricing, micro payments, and subscription models. Regardless of the pricing model, however, managers are understandably concerned about the performance impact of the move from free to fee. Consumers have been conditioned to expect free online content, and there often are many, sometimes thousands of, competing alternatives from which they can choose. When the *New York Times* moved to a paid version of its editorial pages, one reader commented, "The only thing the [*New York Times*] will achieve is a reduced readership for its writers. No one will, nor should we, pay for online content. It's a big, free Web out there" (Haga 2005). The ultimate abandonment of the payment system (on September 19, 2007) appears to vindicate these arguments and illustrates the challenge content providers face online.

In summary, despite high managerial importance, little empirical research has addressed the following questions that companies face when they move from free to fee:

1. What are the long-term performance effects of such a move? In particular, what are the sources of revenue loss that might offset the revenue gains from user fees?
2. What is the role of marketing actions in the free- and fee-user response? In particular, which actions are especially successful in increasing free and fee subscriptions, for both short-term and long-term contracts?

This research attempts to shed light on these issues by focusing on the adoption behavior for free- and fee-based Web content and marketing actions that may affect this behavior. Indeed, the numbers of new free and paying subscribers are key performance indicators for content providers, which depend on these numbers for both advertising and subscription income. The potential drivers of these subscriber numbers are many, including consumer characteristics (e.g., price sensitivity, perceived content value), competitive characteristics (e.g., market concentration, content overlap, pricing schemes), and company characteristics (e.g., content restriction, marketing decisions). In this article, we are interested in the short- and long-term effects of marketing communications and price promotions (i.e., temporary price cuts) on new free- and fee-user subscriptions.[3] To this end, we apply several modern time-series techniques to a daily data set from a commercial Web site in the "business content" industry. Next to "personals/dating" and "entertainment/lifestyles," this industry is among the top three categories for online content, with a 2005 total spending of $312 million and a growth of 49% compared with 2001 (Online Publishers Association 2005).

Research Background

Previous research has examined several phenomena related to our research questions. Three decades ago, Scott (1976) studied the effect of a free two-week trial subscription on consumer willingness to pay for a six-month newspaper subscription. Notably, consumers who were offered the free trial were less likely than those in the control group to become paying subscribers. More recently, Gedenk and Neslin (1999) and Bawa and Shoemaker (2004) analyzed the impact of free samples on subsequent (paid) purchases. The former study found that distribution of free samples increased the probability that consumers would buy the brand in the future. The latter study explicitly distinguished and allowed for potential negative effects (cannibalization of paid purchases) and positive effects (acceleration of purchases by customers and expansion; i.e., purchases by noncustomers). Importantly, both short-term and long-term effects of free samples were positive; they yielded higher customer retention after trial, a larger potential for acceleration of purchases by new consumers, and a higher purchase probability among those who would otherwise not have tried the brand.

Although our research shares a similar focus on the performance effects of marketing actions, it differs from previous studies in at least two important ways. First, all the studies we mentioned examine the effects of marketing actions on purchase behavior in contexts in which people are used to paying, and ultimately expect to pay, for the product or service. In contrast, we examine markets in which (potential) customers have likely become accustomed to free content and thus have a zero reference price. It is not clear how this will affect the impact of marketing on subscriptions. Therefore, our exploratory analysis may provide new insights into marketing effects in such an environment. Second, the move from free to fee is virtually always partial for online content. Firms maintain a mix of free and for-fee content and, therefore, of free and for-fee subscribers. Because content providers often rely on growing free users to attract advertisers, this research is unique in examining how marketing decisions affect both paying customers (who yield subscription revenue) and nonpaying customers (who yield advertising revenue).

In our context, the revenue implications of the move from free to fee may be written in terms of changes to free subscribers (Fsub) and paying subscribers (Psub):

$$\Delta \text{Revenue} = \Delta \text{Fsub} \times \text{Advertising revenue per free subscriber} + \Delta \text{Psub} \times \text{Fee}, \tag{1}$$

where Δ is the difference between the actual value of a series and its benchmark (i.e., a model-forecasted value of the series in the absence of the move). The first part of Equation 1 represents a potential revenue loss from fewer free subscribers, and the second part represents a potential revenue gain from paying subscribers. We proceed by considering the likely drivers of this revenue loss and revenue gain.

Sources of Revenue Loss: How Does the Move from Free to Fee Hurt Free-User Growth?

In a general sense, content providers that move from free to fee are making two strategic decisions: a restriction-level decision and a pricing decision.[4] The restriction-level decision is the extent to which formerly free content is restricted to paying users. Conceptually, restriction can range along a continuum, from complete restriction to highly selective restrictions (i.e., most of the content is free, except for a few selections). Often, at least part of the content remains free (e.g., the *New York Times*, ESPN, our data provider), resulting in a combination of free and fee subscribers.

---

[3] These marketing actions are variables in our data set. In general, content providers use many tactics to encourage payment for content, including free daily passes and free trial subscriptions to protected content.

[4] Data limitations prohibit us from analyzing the effect of these strategic marketing decisions, leaving us with the tactical marketing actions, such as price promotions and e-mails.

When content providers decide to move from free to fee, their site may become less attractive to free users. There are at least two reasons for this. First, such change blocks free access to certain aspects of the provider's content, thus reducing free users' perceived benefits and, consequently, demand (Brynjolfsson, Hu, and Smith 2003). Second, potential users are typically accustomed to obtaining content for free on the Web and likely expect the same from any specific provider. Thus, any (short-term or long-term contract) price is likely to be rejected (Brehm 1966) by some potential users whose preferences for free content have been formed. Fitzsimons and Lehmann (2004) note that encountering unexpected paid content tends to decrease the attractiveness of free content and to reduce satisfaction with the choice options. Conversely, a move from free to fee may expand the market for fee content because of price signaling—that is, the price of fee content signals a higher quality to the (potential) subscriber (Zeithaml 1988).

In addition to its direct effect on new free subscribers, the move from free to fee may decrease the effectiveness of marketing communications on generating free subscriptions. Indeed, the short-term and long-term effects of marketing communications likely depend on the perceived value of the offer (Hanssens and Ouyang 2001). By closing off the for-fee part of content to free subscribers, providers may lower the perceived value of a free subscription. Especially harmful would be a refocus of the communication message on stimulating for-fee, rather than free, subscriptions. Such "conversion" efforts typically highlight the benefits of fee content, at the expense of free content.

In summary, the move from free to fee may influence new free subscriptions (1) directly, (2) through changes in the effectiveness of marketing communications, and (3) through conversion efforts, as represented in Equation 2:

(2) $\Delta\text{Fsub} = \Delta\text{Move} + \Delta\text{Marcom effectiveness} + \Delta\text{Conversion}.$

Sources of Revenue Gain: How Do Marketing Actions Stimulate Fee-User Growth?

Given this potential for revenue loss, it is crucial for the company's performance to generate paid subscriptions. Typically, online content providers offer the choice between short-term contracts and long-term contracts. Across all online content categories, subscriptions represent 78.4% of content purchases (Online Publishers Association 2005); monthly and annual contracts account for 92% of these subscriptions (the remainder are weekly, quarterly, and semiannual forms of subscriptions). The key distinction between short-term and long-term contracts is the length of time the user commits to paying for the content. Commitment issues are particularly important when firms move from free to fee; users accustomed to the no-commitment, no-trade-off situation of free content (Ariely and Shampan'er 2004) suddenly face the choice of whether, and for how long, to commit. We expect that price promotions and marketing communications have different effectiveness in stimulating new short-term versus long-term contract subscriptions.

Some users will be reluctant to make a long-term commitment. Short-term contracts allow such users to try a product with minimal commitment and to diagnose the extent to which the content is useful, thus facilitating their adoption decision (Rogers and Shoemaker 1971). Regret is less likely to occur with a short-term contract; if users find better or equivalent content at a lower price, they can switch sooner with a short-term contract than with a long-term contract. To the extent that price promotions offer a risk premium for trial (Gedenk and Neslin 1999; Pauwels, Hanssens, and Siddarth 2002), they may be effective for stimulating short-term contracts.

Price promotions may be less effective in stimulating long-term contracts. Users willing to commit to such contracts believe that the content will interest them over a long period (Laurent and Kapferer 1985). Moreover, when accessing the content is considered a virtue (versus a vice), users may prefer a long-term contract for precommitment reasons (Wertenbroch 1998). Finally, users may prefer long-term contracts for convenience reasons; that is, they can renew their subscriptions less often (Lambrecht and Skiera 2006). This factor appears to be especially powerful when someone else (e.g., the user's employer) is paying the bill. Given these reasons, price promotions should be less powerful in increasing new long-term contracts.

Instead, marketing communications may play a larger role in generating new long-term contracts. Among such actions, content-based Web sites spend most communication dollars to obtain a high ranking on search-engine sites, such as Google, Yahoo, and MSN, and invest in the optimization of a site so that Web users can find it during an Internet search (Telang, Boatwright, and Mukhopadhyay 2004).[5] Moreover, many sites employ targeted e-mails to free subscribers, with the intent of upselling them to paying services. In summary, we consider how moving from a free- to a restricted fee-based offering affects (1) the "advertising revenue" component of the online content provider, through free-user growth; (2) the "subscription revenue" component, through paid-user growth; and (3) the effectiveness of marketing actions to stimulate free- and paid-user growth.

Empirical Setting

Our data set covers a period of almost four years (June 14, 2001–March 27, 2005) with 1383 daily observations at an online content provider. Its site permits visitors to read articles that provide analysis and opinion on topics of interest to marketing professionals. Visitors register for free content by providing an e-mail address and choosing a password.[6]

---

[5] Other marketing communication tools are banners of various types on external sites, as well as the use of unsold inventory on the focal site.

[6] In principle, free users can be anybody who reads content for free. Unfortunately, we were not able to obtain the complete log files or to distinguish who in the existing log files were paid users. Moreover, we do not know the renewal rates of the paid sub-

The site was supported solely through advertising income until January 1, 2004, at which time a fee-based subscription business model was introduced in which paying subscribers received access to advanced articles and other services, such as a buyer's guide and the guarantee that they would see no third-party advertising.[7] Because advanced articles used to be free, this situation likely represented a loss of value to the free subscribers. Paying subscribers were given the option of either a monthly or a yearly fee, corresponding to our distinction between short-term and long-term contracts. Promotions on these fees were run periodically. Because the restriction of free content occurred on the same day as the charging of fees for accessing the restricted content (as is usual for content providers), we cannot separate these two effects. Instead, we refer to January 1, 2004, as the move from free to fee. Thus, our data capture a single "natural experiment," the analysis of which has a long tradition in the social sciences, including marketing (e.g., Ailawadi, Lehmann, and Neslin 2001; Neslin and Shoemaker 1983). For this proprietary data set, Table 1 displays summary statistics on the subscription and price series.

On average, 111 users each day register for free content, and this number is growing by approximately 5.5 per week (.185 a day). As for paid subscriptions, yearly contracts are more popular than monthly contracts, probably because the former presents a better deal for long-term use. Indeed, the regular monthly price is $4.95, whereas the regular yearly price is $49.95 throughout the sample, and several price promotions are run (in the model, we use these price promotions as variables because the charging of regular prices occurs at the same time as the move from free to fee). These prices are representative of the entire online content industry, for which the latest available figures indicate an average annual subscription fee of $49.69 in 2001 and $48.94 in 2002 (Online Publishers Association 2003).

As for the content provider's marketing efforts, we have data on two types: targeted communication through e-mails to current subscribers and mass-market communication yielding search-engine referrals. Targeted communication includes e-mail "blasts" sent to potential and current free subscribers, with the intention of upselling them to paid-subscription levels. Moreover, we have data on other e-mails sent to a certain number of subscribers, with a variety of news, including announcements of new services and events. In terms of mass-market communication, our data include the number of referral searches from all major search engines to the Web site.[8] This variable is not truly a marketing control variable but rather the outcome of consumer searches and a host of contributing management activities, such as search-engine optimization (which the company started about a year before the move), links from other sites, and a multitude of other promotion and advertising activities that a site engages in to increase its rankings. Given the absence of information on the timing and magnitude of each of these activities, search-engine referrals represent our best proxy measure. For these variables, Table 2 contrasts the average (daily) levels before and after the move from free to fee.

Search-engine referrals increase from 410 in the for-free period to 2921 in the for-fee period. Likewise, e-mails increase from 505 to 7085. In other words, both activities are higher in the period after the move from free to fee and during the for-fee period when virtually all blasts aiming to upgrade (potential) free subscribers to the paid service occur (the one blast in the for-free period occurred shortly before the move, and we do not include it in our analysis, because our series of paying subscriptions starts on January 1, 2004). Finally, we compare the average of daily new free subscriptions before and after the move; this number

TABLE 1
Descriptive Statistics of Daily New Subscriptions and Subscription Prices

	M	Mode	SD	Trend[a]
New free subscriptions	111	82	96.6	.185
New monthly subscriptions	2.06	1	3.24	.0125
New yearly subscriptions	5.55	3	6.85	.0181
Monthly subscription price	$4.94	$4.95	$.062	–.0001
Yearly subscription price	$48.0	$49.95	$5.30	–.0131

[a] Significant trend estimate when we control for daily seasonality, intercept, and autoregression (AR[1]).

TABLE 2
Comparison of Free- Versus Fee-Period Daily Subscriptions and Marketing Actions

	Free-Period Mean	Fee-Period Mean
Search-engine referrals	410.26	2921.32
Targeted e-mail offers	505.18	7084.83
Blasts to switch to fee	1 time	29 times
New free subscriptions	64.13	208.57

---

scribers. Free users typically stay on (they are unlikely to remove themselves from the list) but often become less active over time. Therefore, the data provider communicated that new registrations of free and fee users were its primary performance criteria.

[7] This lack of advertising income from paid subscribers enables us to calculate the trade-off between free-subscriber loss and paid-subscriber gain. Other reasons to keep adding free subscribers (e.g., the creation and maintenance of critical mass for community building) are harder to quantify and are not considered in the analysis. Other content providers may make different decisions, such as continuing third-party advertising to paid subscribers. We leave the analysis of such scenarios for further research.

[8] The company does not engage in paid advertising, because it believes that natural search (i.e., through search-engine optimization) is the best way to attract people interested in the site. This appears to be typical for (smaller) online content providers, in contrast to several companies that sell physical products online.

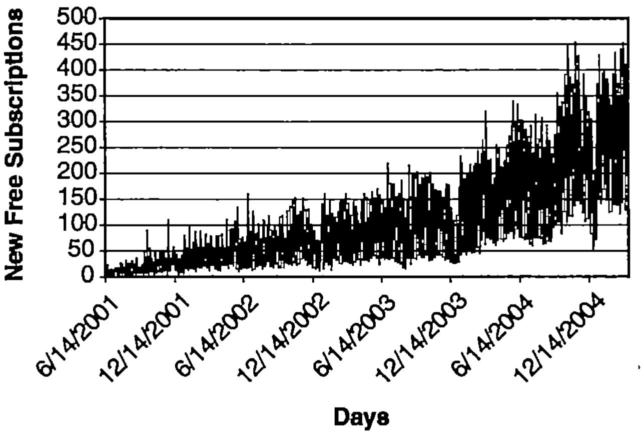
increased from 64 to 209. This growth is also evident from Figure 1, which displays the free subscription time series.

Apart from the positive trend, a key feature of the new subscription data is seasonality of two types: high frequency (day of week) and low frequency (year-end holiday). Pauwels and Dans (2001) observe similar patterns in daily data for online newspapers and offer several reasons. First, the Web site is dedicated to the marketing profession, so new users typically discover it during a workday, not the weekend or holidays. Second, Internet access is typically available at work but may not be available at home (especially in developing countries, from which our data provider attracts some subscribers). Note also that as in Table 2, Figure 1 does not show any easily recognizable harm of the move from free to fee (January 1, 2004) to new free subscriptions. It could be that no such harm occurred or that it is masked by the content provider's increased marketing efforts. Therefore, to address our research questions, we need to go beyond a before/after comparison and formulate a model that accounts for marketing actions.

Methodology

In light of our research questions and data, our choice of methodology is driven by four main considerations. First, the model should control for trend and seasonality patterns, and it must provide a forecasted, expected baseline for each performance variable so that we can capture the impact of unexpected events (e.g., the move from free to fee) as deviations from this baseline (Abraham and Lodish 1987). Second, the model should be robust to deviations from stationarity, which can lead to spurious regression problems (Granger and Newbold 1986).[9] Third, the model should provide a flexible treatment of both short-term and long-term effects of marketing actions on performance (Dekimpe and Hanssens 1999). Fourth, the model should allow for various dynamic interactions among performance variables (e.g., the likely interaction between new free and paid subscriptions), among marketing variables (e.g., a likely increase in marketing communication each time a fee price is promoted), and between marketing and performance variables (e.g., a decrease in new free subscriptions induces the content provider to send more e-mails).[10]

FIGURE 1
Daily Data for New Free Subscriptions



In summary, our model should account for the time-series properties of performance and marketing variables and for their dynamic interactions. The persistence modeling framework (Dekimpe and Hanssens 2004) addresses these four criteria in several methodological steps, which we summarize and extend (with robustness checks) in Table 3.

Permanent Versus Temporary Change: Unit Root and Cointegration Tests

Unit root tests verify the univariate time-series properties (stationarity versus evolution) for each variable. The substantive question they address is whether our performance variables (new subscriptions) are mean reverting (stationarity) or have changed permanently in the data sample (evolution). We use both the augmented Dickey–Fuller test procedure (see Enders 2003) and the KPSS test (Kwiatkowski et al. 1992). The former maintains evolution as the null hypothesis (and is the most popular in marketing applications), and the latter maintains stationarity as the null hypothesis. Each test is estimated in two forms: with and without a deterministic time trend. Convergent conclusions of these different tests yield greater confidence in our variable classification (Maddala and Kim 1998). Next, we examine the robustness of our unit root test conclusion to structural breaks (Perron 1989, 1990) because the move from free to fee is an obvious choice for such a break point. Finally, Johansen, Mosconi, and Nielsen's (2000) cointegration test verifies whether any combination of evolving variables is in long-term equilibrium, allowing for structural breaks in these variables.

Modeling Dynamic Interactions: Vector-Autoregressive Models

Next, we specify vector-autoregressive (VARX) models that are well suited to measuring the dynamic performance response and interactions between performance and marketing variables (Dekimpe and Hanssens 1999). Both performance variables and marketing actions are endogenous; that is, they are explained by their own past and the past of the other endogenous variables. Specifically, VARX models not

---

[9] Stationary variables fluctuate as temporary deviations around a fixed mean or trend. Evolving variables have a unit root; that is, they fluctuate without reversion to a fixed mean or trend. For technical definitions and applications in marketing, see Dekimpe and Hanssens (1999) and Pauwels and colleagues (2004).

[10] We note that the restricted and free content varies over time as new (restricted and free) articles are added to the site. Therefore, at any given day, the perceived consumer benefits of a free subscription versus a paid subscription likely depend on the perceived value of the restricted versus free content. In other words, we expect that there is a negative correlation between free subscriptions and paid subscriptions, which logically leads to a dynamic system model, such as a VARX model or the simultaneous equation system we estimate.

TABLE 3
Overview of Methodological Steps

Methodological Step	Relevant Literature	Research Question
1. **Unit Root and Cointegration Tests**		
Augmented Dickey–Fuller test	Enders (2003)	Are variables stationary or evolving?
KPSS test	Maddala and Kim (1998)	Are the unit root results robust to null hypothesis?
Perron test	Perron (1989, 1990)	Are the unit root results robust to structural breaks?
Cointegration test	Johansen, Mosconi, and Nielsen (2000)	Are evolving variables in long-term equilibrium?
2. **Model of Dynamic Interactions**		
VARX model	Pauwels, Hanssens, and Siddarth (2002) Pauwels et al. (2004)	How do sales and marketing variables interact in the long run and the short run, when the unit root and cointegration results are taken into account?
VARX in differences	Dekimpe and Hanssens (1999)	
Vector error correction model	Dekimpe and Hanssens (2004)	
3. **Policy Simulation Analysis**		
Impulse-response function	Dekimpe and Hanssens (1999)	What is the dynamic (performance) response to a (marketing) impulse?
Generalized impulse response	Pesaran and Shin (1998)	What is the immediate effect of an impulse, without imposing a causal ordering?
Long-term marketing effect	Dekimpe and Hanssens (2004)	What is the long-term impact of a marketing impulse on performance?
4. **Robustness Checks**		
Bootstrapping model parameters	Bradley and Tibshirani (1993)	How stable are the model parameters?
Estimation of regression and simultaneous equation model	Zellner and Theil (1962)	Do the VARX models outperform simpler models with dynamic marketing effects on performance?

only measure short-term and long-term response to marketing actions but also capture the performance implications of complex feedback loops.

Vector-autoregressive models are specified in levels or differences (d) of each endogenous variable, depending on the results of its unit root tests. Model specification requires two remaining choices: determining the number of lags K, also known as the order of the model, and determining which variables to include as endogenous. We base the former on the Bayesian information criterion, which is a consistent estimator of lag length (Lütkepohl 1993), and test whether we should add lags to pass diagnostic tests on residual autocorrelation (Franses 2005). Regarding the latter, we include all variables as endogenous for which Granger-causality tests reveal that they are caused by another variable in the system (Enders 2003). The remaining variables (in our case, only the move from free to fee) are included as exogenous variables, with as many lags as indicated by residual tests.[11] Finally, as control variables (the vector C), we include an intercept, six day-of-week seasonal dummies (in which Friday is the benchmark), a dummy for the year-end holidays, and a time trend to capture external factors (to the extent possible), such as growth in Internet access, growth in people with high-speed bandwidth, and general increases in content and content providers.

Given the nature of our data and research focus, we must be careful which data periods to analyze. There were no paying subscribers before January 1, 2004, and the regular price jump from zero to the regular (monthly and yearly) prices is likely to have had a different impact than the subsequent price promotions off the regular price. Therefore, we specify two different models.

Model 1. For the full data period, Model 1 includes new free subscriptions, both marketing efforts (e-mails and blasts), and search-engine referrals. The impact of the move from free to fee on free subscriptions manifests in two ways: (1) The pulse variable "move" (taking on a value of 1 at the time of the move from free to fee) measures the direct effect on the subscription series, and (2) the step variable "fee" (0 in the for-free period and 1 in the fee-and-free period) interacts with search-engine referrals and e-mail to assess how their effectiveness differs after the move.[12] All variables may have an immediate (same-day) or a lagged effect. Equation 3 presents the model:

---

[11] All Granger-causality tests with lags of up to 120 indicate that new free subscribers do not cause the move from free to fee. As a robustness check, we included the move from free to fee as an endogenous variable, and indeed we found no difference in our substantive results.

[12] We do not need to interact blasts with fee, because blasts are logically related to the fee period (see the "Empirical Setting" section).

$$(3)\quad \begin{bmatrix} \text{dFsub}_t \\ \text{dSearch}_t \\ \text{E-mail}_t \\ \text{Blast}_t \\ \text{Search} \times \text{Fee}_t \\ \text{E-mail} \times \text{Fee}_t \end{bmatrix} = \text{C}_1 + \sum_{k=1}^{K} \text{B}_k \times \begin{bmatrix} \text{dFsub}_{t-k} \\ \text{dSearch}_{t-k} \\ \text{E-mail}_{t-k} \\ \text{Blast}_{t-k} \\ \text{Search} \times \text{Fee}_{t-k} \\ \text{E-mail} \times \text{Fee}_{t-k} \end{bmatrix}$$

$$+ \sum_{j=0}^{J} \Gamma_j \times \begin{bmatrix} \text{Move}_{t-j} \\ \text{Move}_{t-j} \\ \text{Move}_{t-j} \\ \text{Move}_{t-j} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} u_{\text{Fsub},t} \\ u_{\text{Search},t} \\ u_{\text{E-mail},t} \\ u_{\text{Blast},t} \\ u_{\text{S} \times \text{F},t} \\ u_{\text{E} \times \text{F},t} \end{bmatrix},$$

where Fsub is the number of new free subscriptions, Search is the number of search-engine referrals, E-mail is the number of general e-mails sent, Blast is the presence of free-to-fee conversion e-mails, C is the (6×1) vector of control variables, K is the order of the model, B_k and Γ_j are the (6×6) and (6×1) vectors of dynamic coefficients, J is the maximum of lags selected for the effect of move on each endogenous variable, and $(u_{\text{Fsub},t}, \ldots, u_{\text{E} \times \text{F},t})' \sim N(0, \Sigma_u)$. We estimate this model in units (instead of logarithms) to obtain a managerially relevant decomposition of the effects of interest (Van Heerde, Gupta, and Wittink 2003).

Model 2. For the for-fee period starting January 1, 2004, Model 2 includes new free subscriptions and all three marketing efforts (e-mails, search, and blasts) and adds new subscriptions for the monthly (Msub) and yearly (Ysub) contracts. Moreover, instead of the regular prices (which are constant throughout the for-fee period), we include price promotions (dollars off) for monthly (Mprom) and yearly (Yprom) contracts:

$$(4)\quad \begin{bmatrix} \text{dFsub}_t \\ \text{dMsub}_t \\ \text{dYsub}_t \\ \text{dSearch}_t \\ \text{E-mail}_t \\ \text{Blast}_t \\ \text{Mprom}_t \\ \text{Yprom}_t \end{bmatrix} = \text{C}_2 + \sum_{k=1}^{K} \text{B}_k \times \begin{bmatrix} \text{dFsub}_{t-k} \\ \text{dMsub}_{t-k} \\ \text{dYsub}_{t-k} \\ \text{dSearch}_{t-k} \\ \text{E-mail}_{t-k} \\ \text{Blast}_{t-k} \\ \text{Mprom}_{t-k} \\ \text{Yprom}_{t-k} \end{bmatrix} + \begin{bmatrix} u_{\text{Fsub},t} \\ u_{\text{Msub},t} \\ u_{\text{Ysub},t} \\ u_{\text{Search},t} \\ u_{\text{E-mail},t} \\ u_{\text{Blast},t} \\ u_{\text{Mprom},t} \\ u_{\text{Yprom},t} \end{bmatrix},$$

where Fsub is the number of new free subscriptions, Msub is the number of new monthly contract subscriptions, Ysub is the number of new yearly contract subscriptions, Search is the number of search-engine referrals, E-mail is the number of general e-mails sent, Blast is the presence of free-to-fee conversion e-mails, Mprom is the dollar amount of a price promotion on the monthly subscription price, Yprom is the dollar amount of a price promotion on the yearly subscription price, C is the (8×1) vector of control variables, K is the order of the model, B_k is the (8×8) vector of dynamic coefficients, and $(u_{\text{Fsub},t}, \ldots, u_{\text{Yprom},t})' \sim N(0, \Sigma_u)$. We estimate Equation 4 both in units and in logarithms. The former yields a managerially relevant assessment of marketing effectiveness, and the latter directly yields unit-free elasticities (Pauwels, Hanssens, and Siddarth 2002), which allow for the comparison of percentage gains of marketing communication and price promotions for monthly versus annual contracts.

Long-Term Impact of Marketing Actions: Impulse-Response Functions

A VARX model estimates the baseline of each endogenous variable and forecasts its future values on the basis of the dynamic interactions of all jointly endogenous variables. Based on the VARX coefficients, impulse-response functions track the over-time impact of unexpected changes (shocks) to the marketing variables on forecast deviations from baseline for the other endogenous variables.

To derive the impulse-response functions of a marketing action, we compute two forecasts, one based on the information set without the marketing action and the other based on the extended information set that accounts for the marketing action. The difference between these forecasts measures the incremental effect of the marketing action. These dynamic effects are not a priori restricted in time, sign, or magnitude. As for the immediate (same-day) effects, we adopt the generalized, simultaneous-shocking approach (Pesaran and Shin 1998), which uses information in the residual variance–covariance matrix of the VARX model rather than requiring the researcher to impose a causal ordering among the endogenous variables (Dekimpe and Hanssens 1999). Finally, we follow established practice in marketing research and assess the statistical significance of each impulse-response value by applying a one-standard-error band, as Pesaran, Pierse, and Lee (1993) and Sims and Zha (1999) recommend. Our interpretation of the estimated effects focuses on the immediate (same-day) effect and the permanent effect (i.e., the value at which the impulse-response function stabilizes). These numbers operationalize, respectively, the short-term and the long-term effects of variables on the subscription series. Finally, we compare the long-term effects on the basis of the significance tests for the difference in impulse-response functions, as Pauwels (2004) outlines. In particular, two impulse-response functions are considered significantly different if their long-term effect is significantly different (using the standard error of the accumulated response function, based on the joint distribution of the individual period's standard errors, as detailed in Lütkepohl [1993, Chap. 3.7]).

Results

Unit Root Tests and VARX Model Fit Assessment

The unit root tests (Table 4) reveal evolution in all three subscription variables and in search-engine referrals (even after we allow for a structural break at the time of the move) but reveal stationarity for e-mails, blasts, monthly contract, and yearly contract price promotions. Because no combination of evolving variables shows evidence of cointegration, we proceed by estimating our models with the three sub-

TABLE 4
Summary of the Unit Root Test Results

	Augmented Dickey–Fuller Test Without Trend	KPSS Test Without Trend	Augmented Dickey–Fuller Test with Trend	KPSS Test with Trend
Free subscriptions	*–.23*	*4.02*	*–2.87*	*.67*
Monthly contract subscriptions	*–1.20*	*2.73*	*–3.18*	*.19*
Yearly contract subscriptions	*–3.10*	*1.33*	–4.10	*.19*
Search-engine referrals	*–.38*	*3.88*	*–2.58*	*.87*
E-mails	–6.24	42	–33.93	.06
Blasts	–22.47	.32	–22.56	.05
Monthly contract price promotion	–5.04	.33	–5.27	.08
Yearly contract price promotion	–4.88	.37	–6.68	.12

Notes: Number in italics indicate that the series is classified as evolving.

scription variables and search-engine referrals in differences. Significant marketing effects on performance difference (or "growth") imply that the action has permanently changed the performance level (Dekimpe and Hanssens 1999; Pauwels et al. 2004).

Both estimated VARX models have significant F-statistics for overall model fit (95.5 and 14.65, respectively). Moreover, they adequately explain the performance series. The adjusted R-squares for new free-subscription growth are .832 for Model 1 and .865 for Model 2; for new monthly and yearly subscription growth, the adjusted R-squares are, respectively, .597 and .683 (Model 2). Likewise, growth in search-engine referrals follows predictable patterns. The explanatory power for the marketing action variables is logically lower (between .12 and .28), indicating that they are mostly unexpected "shocks" to the system and, thus, plausibly unexpected to customers as well. Finally, both models pass the diagnostic tests (Franses 2005) for residual correlation (the Durbin–Watson test and Lagrange-multiplier tests), residual normality (the Jarque–Bera test), and heteroskedasticity (the White test).

Short-Term and Long-Term Marketing Effects on New Subscriptions

Estimation of the impulse-response functions reveals that all dynamic effects stabilize within a week, as Figures 2–4 illustrate for the main driver affecting each of the three new subscription series.[13] The impact of search-engine referrals on new free subscriptions is significant in the first day, peaks in the second, and then settles into a significant long-term effect (Figure 2). In contrast, monthly contract price promotions have their greatest impact right away on new monthly contract subscriptions and then show a postpromotion dip before stabilizing at a significant, positive long-term level (Figure 3). In comparison, the effect of yearly contract price promotions on new yearly contract subscriptions is much smaller in magnitude. However, a small post-promotion dip is offset by positive dynamic effects, causing the long-term effect to exceed the short-term effect. Plausibly, yearly contract subscribers are not as price sensitive as

FIGURE 2
Impulse-Response Function of New Free Subscriptions to Search-Engine Referrals

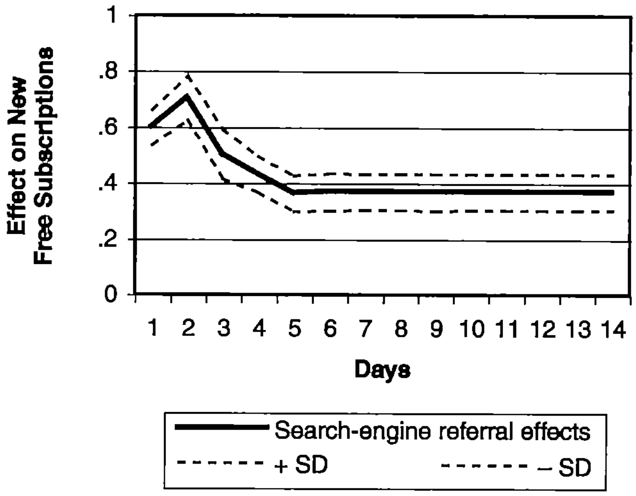
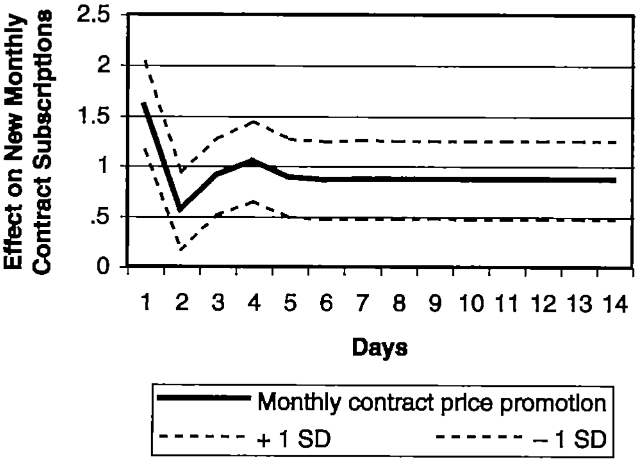


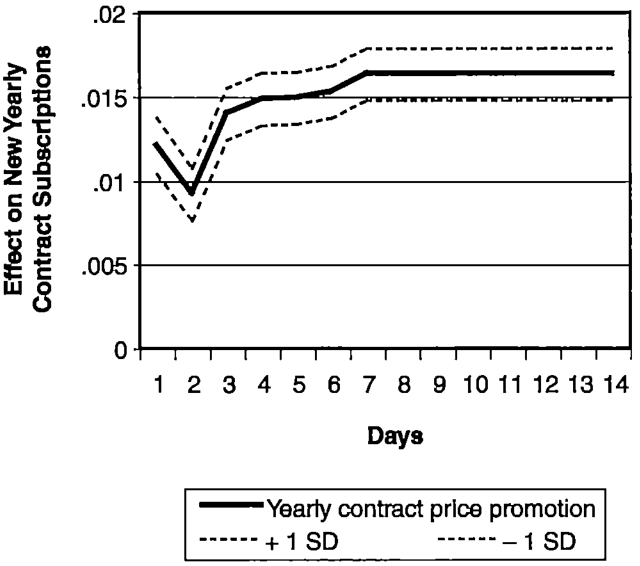
FIGURE 3
Impulse-Response Function of Monthly Contract New Subscriptions to Monthly Contract Price Promotions



---

[13] Because all subscriber series enter the model in differences, we calculate and display the accumulated impulse-response functions to obtain the marketing effects on the levels of each subscriber series.

FIGURE 4
Impulse-Response Function of Yearly Contract New Subscriptions to Yearly Contract Price Promotions



monthly contract subscribers, and the promotion mostly acts as a salient attention grabber, which leads several people to sign up a few days afterward. Tables 5–6 present the full results on short- and long-term marketing effects.

Table 5 displays the unit effects of potential drivers on new free subscriptions. Both search-engine referrals and e-mails have a positive effect on new free subscriptions. However, the move from free to fee reduces the growth in free subscriptions compared with the baseline forecast. This reduction is relatively substantial; the daily number of new free subscriptions is 208 fewer in the for-fee period than it would have been had the company not started charging for content.

Moreover, the move from free to fee reduces marketing communication effectiveness. Both search-engine referrals and e-mails have lower long-term effects on new free subscriptions: 6% (–.029/.453) and 21% (–.525/2.466) fewer than before the move. We can multiply this loss in marketing communication effectiveness by the average marketing communication levels in Table 2 to conclude that the company lost 79 daily free subscriptions through the search-engine referral effect and 4 daily free subscriptions through the e-mail effect.

Finally, there is a notable twist in the effect of blasts on new free subscriptions. The short-term effect is positive, indicating a possible complementarity effect; namely, the general e-mail blasts may further raise awareness for the free site and induce free customers to recommend the site to friends and colleagues. In the long run, however, the substitution effect dominates; the blast reduces free-subscriber growth, probably because it convinces potential free subscribers to become paid subscribers instead. Indeed, Table 6 shows that blasts increase both monthly contract and yearly contract subscriptions (we provide a closer examination of this trade-off subsequently).

The marketing communication and promotion effects in Table 6 show the expected signs, including some cannibalization of monthly subscriptions through promotions on the

TABLE 5
Short-Term and Long-Term Effects of Marketing Actions on New Free Subscriptions

	VARX Model 1		VARX Model 2	
	Short-Term	Long-Term	Short-Term	Long-Term
Search engine	.866	.453	.597	.366
Search engine × fee	–.016	–.029		
E-mails (× 1000)	5.244	2.466	4.853	1.489
E-mails (× 1000) × fee	–.272	–.525		
Blast	60.742	–25.917	77.151	–38.831
Move from free to fee	–48.901	–208.059		
Monthly price promotion			–4.358	–7.684
Yearly price promotion			.000	–.999

TABLE 6
Short-Term and Long-Term Effects of Marketing Actions on New Paid Subscriptions

	Monthly Contract Subscriptions		Yearly Contract Subscriptions	
	Short-Term	Long-Term	Short-Term	Long-Term
Search engine	0	.001	0	.004
E-mails (× 1000)	0	.055	.252	.119
Blast	.194	.173	1.092	1.745
Monthly price promotion	1.589	.858	0	0
Yearly price promotion	0	–.003	.012	.016

Notes. All nonzero effects are significant; nonsignificant effects are displayed as 0.

yearly contract's price. Notably, monthly price promotions have a stronger long-term unit effect on monthly subscriptions than yearly price promotions have on yearly subscriptions. As for communication efforts, all three variables have a greater long-term unit impact on new yearly subscriptions than on new monthly subscriptions. E-mails are twice as effective in stimulating yearly versus monthly subscriptions, and search-engine referrals are three times as effective in generating new yearly versus monthly subscriptions. Finally, general e-mail blasts (designed to induce free subscribers to upgrade) are ten times more effective for stimulating yearly contracts than for stimulating monthly contracts. These long-term effect differences are significant at the 1% level.

Table 7 shows the estimated marketing elasticities based on the log-log version of Equation 4. Although we have no such comparison benchmark for the other marketing actions, the promotional elasticity for monthly contracts is in the range reported for sales of fast-moving consumer goods (e.g., Pauwels 2004, 2007), and the promotional elasticity for the yearly contracts is in the range reported for newspaper subscriptions (Lewis 1995) and access to telecommunications services (Danaher 2002). We observe the same relative pattern as that in Table 6; namely, price promotions have a greater long-term elasticity for monthly contracts, and marketing communication efforts have a greater long-term elasticity for yearly contracts (elasticity differences are significant at the 5% level). However, the magnitude of the difference is less outspoken because yearly contracts have a higher baseline than monthly contracts. Despite the inelastic demand for the yearly contract, it receives most of the promotional activity. This may indicate that management overestimates the price elasticity of such subscriptions, as many scholars report for the printed newspaper industry (Lewis 1995).

Robustness Checks

We assess the robustness of our substantive results in two ways. First, we investigate the parameter stability of the VARX models by bootstrapping. We draw a sample with replacement 250 times and estimate the model on each sample. A comparison of the original estimates and the distribution of the bootstrap estimates reveals that all substantive results hold. For each significant effect in the original analysis, we observe that the 90% range of the bootstrapped estimates excludes 0 (detailed results are available on request). Thus, there is no cause for concern regarding the stability of the parameters of interest, because the results are consistent with the information in the original coefficient estimates and their standard errors.

Second, we estimate a regression model explaining dFsub that corresponds to the VARX model in Equation 3 and a simultaneous equation system that corresponds to the VARX model in Equation 4, which we display in Equation 5:

$$
\begin{aligned}
(5)\quad \mathrm{dFsub}_t = {} & \mathrm{C}_{\mathrm{Fsub},t} + \sum_{k=1}^{K} \delta_{\mathrm{F}}^{k} \times \mathrm{dFsub}_{t-k} \\
& + \sum_{l=0}^{L} \gamma_1^{l} \times \mathrm{dMsub}_{t-l} + \sum_{m=0}^{M} \gamma_2^{m} \times \mathrm{dYsub}_{t-l} \\
& + \sum_{n=0}^{N} \beta_1^{n} \times \mathrm{dSearch}_{t-n} + \sum_{p=0}^{P} \beta_2^{p} \times \text{E-mail}_{t-p} \\
& + \sum_{q=0}^{Q} \beta_3^{q} \times \mathrm{Blast}_{t-q} + \sum_{r=0}^{R} \beta_4^{r} \times \mathrm{Mprom}_{t-r} \\
& + \sum_{v=0}^{V} \beta_5^{v} \times \mathrm{Yprom}_{t-v} + \mathrm{u}_{\mathrm{Fsub},t}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{dMsub}_t = {} & \mathrm{C}_{\mathrm{Msub},t} + \sum_{k=1}^{K} \delta_{\mathrm{M}}^{k} \times \mathrm{dMsub}_{t-k} \\
& + \sum_{l=0}^{L} \gamma_3^{l} \times \mathrm{dFsub}_{t-l} + \sum_{m=0}^{M} \gamma_4^{m} \times \mathrm{dYsub}_{t-l} \\
& + \sum_{n=0}^{N} \beta_6^{n} \times \mathrm{dSearch}_{t-n} + \sum_{p=0}^{P} \beta_7^{p} \times \text{E-mail}_{t-p} \\
& + \sum_{q=0}^{Q} \beta_8^{q} \times \mathrm{Blast}_{t-q} + \sum_{r=0}^{R} \beta_9^{r} \times \mathrm{Mprom}_{t-r} \\
& + \sum_{v=0}^{V} \beta_{10}^{v} \times \mathrm{Yprom}_{t-v} \\
& + \mathrm{u}_{\mathrm{Msub},t}
\end{aligned}
$$

TABLE 7
Marketing Elasticity of New Paid Subscriptions (Log-Log Model)

	Monthly Contract Subscriptions		Yearly Contract Subscriptions	
	Short-Term	Long-Term	Short-Term	Long-Term
Search engine	0	1.42	0	2.11
E-mails	0	.19	.32	.32
Blast	.01	01	.01	.02
Monthly price promotion	3.86	2.09	0	0
Yearly price promotion	0	–.01	.11	.14

Notes: All nonzero effects are significant; nonsignificant effects are displayed as 0

$$\begin{aligned} \text{dYsub}_t = {} & C_{\text{Ysub},t} + \sum_{k=1}^{K} \delta_Y^k \times \text{dYsub}_{t-k} \\ & + \sum_{l=0}^{L} \gamma_5^l \times \text{dFsub}_{t-l} + \sum_{m=0}^{M} \gamma_6^m \times \text{dMsub}_{t-l} \\ & + \sum_{n=0}^{N} \beta_{11}^n \times \text{dSearch}_{t-n} + \sum_{p=0}^{P} \beta_{12}^p \times \text{E-mail}_{t-p} \\ & + \sum_{q=0}^{Q} \beta_{13}^q \times \text{Blast}_{t-q} + \sum_{r=0}^{R} \beta_{14}^r \times \text{Mprom}_{t-r} \\ & + \sum_{v=0}^{V} \beta_{15}^v \times \text{Yprom}_{t-v} + u_{\text{Ysub},t}, \end{aligned}$$

where C stands for the deterministic terms (intercept, trend, and seasonality), all variables have the same interpretation as before, and the number of lagged terms (determined by a specification search) may differ for each variable and in each equation. We estimate this model with three-stage least squares, which accounts for the correlation of the right-hand-side variables with the error terms and for contemporaneous correlation and heteroskedasticity in the residuals (Zellner and Theil 1962), using lagged terms as instrumental variables.

Note that the model in Equation 5 controls for trend and seasonality patterns and for deviations from stationarity (variables are included in differences or levels, depending on the unit root test results). Moreover, it incorporates both immediate (short-term) and dynamic (long-term) effects of marketing actions on the subscription series and the possible relationship between free and paid subscriptions. However, it does not incorporate dynamic interactions among the marketing variables or dynamic effects of the subscription variables on the marketing actions ("performance feedback" in Dekimpe and Hanssens [1999]). Therefore, a comparison of model fit for Equations 3 and 4 indicates the extent to which such dynamic interactions help explain the data patterns. We compare both in-sample fit and forecasting accuracy (using the 2005 data as a holdout sample). Finally, a comparison of the model estimates for the common variables provides a robustness check of our substantive results.

Our estimation of the regression and simultaneous equation models yields significant F-statistics for overall fit for each model. We compare these models with the proposed VARX models using (1) in-sample explanatory power and (2) one-step-ahead forecast errors in the holdout sample. Table 8 shows that the VARX models have higher R-squares and adjusted R-squares and lower one-step-ahead root mean square errors than their corresponding equation, for each of the three subscription series in the simultaneous equation model.

Comparing the many model estimates with the corresponding VARX model (results are available on request), we observe a strong consistency in the substantive results (sign and significance) for both models, with an almost identical estimate for the immediate (short-term) effect. The major difference lies in the long-term effects; after the first lag, the effect estimates are virtually all nonsignificant in the alternative models. Plausibly, the VARX model captures performance feedback and other complex dynamic interactions, which have made it the model of choice in many recent studies that focus on long-term effects (for a review, see Dekimpe and Hanssens 2004).

In summary, the robustness checks reveal no reason to doubt the validity of our substantive results, though replication is needed to examine whether they apply to other content providers. We offer several suggestions to this end in the section "Conclusions and Avenues for Further Research."

Decomposing the Sources of Revenue Loss and Comparing Them with Revenue Gains

We assess managerial implications by decomposing the loss of free subscriptions into its sources (Equation 2) and by comparing the resultant loss of potential advertising revenue with paid-subscriber revenue (Equation 1). To this end, we compare the actual number of new free subscriptions with the expected number based on our VARX forecasts.

TABLE 8

Comparison of Fit Statistics for the Simultaneous and VARX Models: In-Sample Explanatory Power and Root Mean Square Forecast Error (RMSE) for Holdout Sample

	Regression Model		VARX Model 1	
	Adjusted R^2	RMSE	Adjusted R^2	RMSE
D(free subscriptions)	.774 (.771)	.628	.840 (.832)	.293
	Simultaneous Equation Model		**VARX Model 2**	
	Adjusted R^2	**RMSE**	**Adjusted R^2**	**RMSE**
D(free subscriptions)	.757 (.753)	.306	.886 (.865)	.272
D(monthly subscriptions)	.531 (.521)	5.924	.661 (.597)	3.653
D(yearly subscriptions)	.502 (.491)	6.631	.737 (.683)	4.557

We obtain the model-forecasted levels of new free subscriptions in the absence of the move from free to fee. This average daily number is 1147 higher than in the for-free period because of the large increase in search-engine referrals and e-mails (Table 2). Because the actual daily number is only 144 higher, we attempt to explain the loss of 1003 free subscriptions a day. Our VARX model estimates reveal the following three drivers:

1. The move from free to fee itself (i.e., the combination of restricting free content and charging for content) accounts for 208 lost subscriptions a day.
2. The decrease in marketing communication effectiveness (search-engine referrals and e-mails) accounts for 71 fewer free subscriptions a day.
3. The blasts in the for-fee period decrease new free subscriptions by 752.

The total estimated loss from our model is 1031, which differs from the raw data loss of 1003 by only 28 (i.e., approximately 3%).

The relative magnitude of the first two revenue loss sources seems intuitive; that is, the reduction in marketing effectiveness matters but not nearly as much as the move from free to fee itself (i.e., the restriction of valuable content that used to be free). The impact of blasts is more surprising, but it makes sense when considered in combination with the gain in new paying subscription numbers. Indeed, although all 29 blasts together decrease new free subscriptions by 752, they increase new monthly contract and yearly contract subscriptions by 5 and 51, respectively. The relevant question then becomes which of the blasts' effects yields the highest monetary impact. In general, we need information on the advertising income generated by free users to compare their revenue contribution with that of the for-fee subscribers.

The content provider gave us its best estimates of how much advertising revenue increases for every 1000 new free subscriptions. This revenue is based almost exclusively on the advertising e-mails sent to the free subscribers; other forms of advertising derive much lower revenues. This additional revenue for an extra free subscriber is $1.63 a year, or $.0045 a day.[14] From Table 1, we know that the comparable revenue numbers are $48 for yearly contract subscribers and $4.94 for monthly subscribers (a conservative estimate, assuming that they do not renew, in which case they are worth up to 12 times more, or $59.28 a year). These paying subscribers are not exposed to e-mail advertising and thus yield only their fee in revenues. We use this number to calculate the net revenue impact of blasts and of the move from free to fee.

The blasts generate $2,454 in paying-subscriber revenue, which is double the total loss of $1,224 in free-subscriber advertising revenue. Overall, the move from free to fee yields a daily revenue loss of $4.64 (1031 × $.0045), which is more than offset by the daily average revenue of $277 (2.06 × $4.94 + 5.55 × $48) from paying subscribers. In other words, the move from free to fee is profitable for this firm; the resultant subscriber revenue is higher than the loss in advertising revenue.

Timing of the Move from Free to Fee

Although we focus on analyzing the consequences of the move from free to fee, its antecedents may also be of interest to managers contemplating such a move. In particular, it seems intuitive to wait until the free site becomes more popular, a sign that consumers have come to realize its value. We test for this intuition in several ways. First, Granger-causality tests determine whether the move can be attributed to daily changes in (new free) subscriptions. Second, the move's timing could also be influenced by trends in subscriptions.[15] Indeed, recent research offers conceptual and empirical support for the importance of such "performance regimes" (Pauwels and Hanssens 2007), operationalized as rolling-window estimates of either the deterministic trend (ibid.) or the stochastic trend in performance (Bronnenberg, Mahajan, and Vanhonacker 2000; Pauwels 2001).

A deterministic trend implies that the series steadily moves in a certain direction. A significant, positive trend represents growth, and a significant, negative trend represents decline. We test for the trend's significance in a rolling-window test that decomposes the performance series into constant, trend, and seasonal fluctuations, as depicted in Equation 6:

$$(6)\qquad y_t = c + \delta t + \sum_{j=2}^{7} \lambda_j \mathrm{DoW}_j + \gamma \mathrm{Holiday} + \rho y_{t-1} + \varepsilon_t,$$

where t represents the deterministic time trend, DoW are the six day-of-week dummies (Friday is the benchmark) (this captures high-frequency noise), Holiday is the dummy variable for end-of-year holidays (this captures low-frequency noise), and y_{t-1} is the lagged dependent variable to capture autoregressive behavior (as indicated by the Durbin–Watson statistic). We estimate Equation 6 in rolling windows of 365 daily observations.[16] This rolling-window analysis (Leeflang et al. 2000; Swanson and White 1997) uses a data sample of fixed size, which allows for direct comparisons between the estimates in different windows. On the basis of these rolling-window tests, Pauwels and Hanssens (2007) classify performance regimes along two dimensions: the performance trend sign (up, nonsignificant, or down) and the change in this trend (accelerating or decelerating). Ranked according to managerial desirability, these regimes are called "accelerating growth" (upward trend with accelerating trend change), "saturating growth" (upward trend with decelerating trend change), "improving stability," "lessening stability," "decline turnaround," and "deteriorating decline."

---

[14] To assess the generalizability of this yearly advertising revenue estimate for a free user, we obtained price quotes from several community Web sites. The average monthly number of pages viewed on a community site by unique visitors is approximately 130. Our observation across these Web sites reveals that the average page carries about two to three advertisements. As a result, the range of yearly advertising revenue from the average (free) user is ($1.25, $1.87), which contains our company's estimate of $1.63.

[15] We thank an anonymous reviewer for this suggestion.

[16] We chose 365 observations for several reasons: (1) This number covers a full year, which enables us to account consistently for seasonality in each window (Pauwels and Hanssens 2007); (2) it represents a sufficiently large sample for each individual window estimation; and (3) it is not so large as to prevent us from picking up transition points around the move from free to fee. Sensitivity analysis reveals that our empirical results hold for windows, varying from 182 to 548 observations (i.e., 6–18 months).

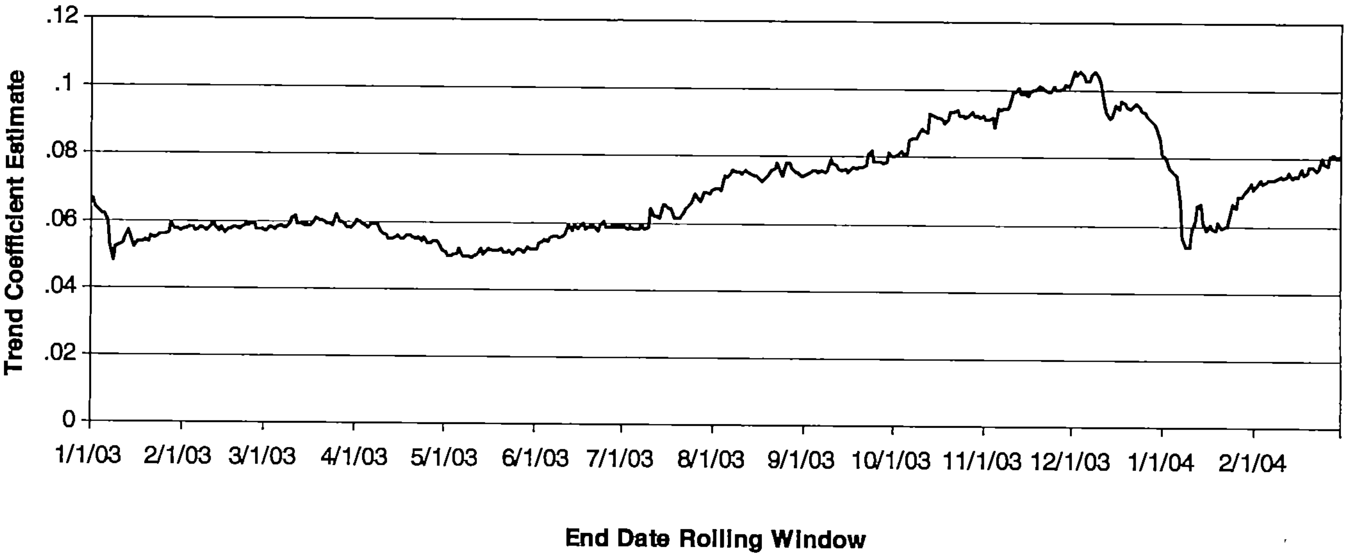
The interpretation of a stochastic trend (evolution) and its absence (stationarity) deserves further explanation. Stationarity, or mean reversion, implies that all observed fluctuations are temporary deviations from deterministic components, which may include mean, trend, and seasonal cycles (as in Equation 6). Thus, if an unexpectedly large number of people sign up today, we would not change our expectation of the number of new subscribers tomorrow. In contrast, evolution implies that no mean reversion occurs; that is, the performance series may wander widely apart from any previously held position. Therefore, if a large number of people sign up today, this "shock" substantially changes our expectation of the number of people who will sign up tomorrow. In the case of such positive evolution, the performance series has momentum (i.e., better performance today generates better performance tomorrow). The augmented Dickey–Fuller test investigates the presence of a stochastic trend as follows:

$$(7)\quad \Delta y_t = \gamma y_{t-1} + \delta t + \alpha + \sum_{i=1}^{p} \beta_i dy_{t-i} + \sum_{j=2}^{7} \lambda_j DoW_j + \gamma Holiday + \varepsilon_t,$$

where dy_t is the first difference of the series and the number of lags, p, selected by the Bayesian information criterion (Hamilton 1994). We add the deterministic time trend to isolate the stochastic trend (Enders 2003), and the seasonal dummies provide good size estimates of the Dickey–Fuller test in seasonal series (Ghysels, Lee, and Noh 1994). The coefficient of interest is γ; the t-statistic testing $\gamma = 0$ is the Dickey–Fuller test for a unit root (evolution) against a trend-stationary alternative. The critical value for this test specification at the 5% significance level is –3.41; t-statistics below this value indicate trend stationarity, and t-statistics above this value indicate evolution. Again, we estimate Equation 6 in rolling windows of 365 daily observations to investigate the stochastic trend across periods and relate it to the timing of the move from free to fee.

The results of our timing analysis are informative. Granger-causality tests provided no evidence that the move from free to fee was caused by daily changes in new free subscriptions, so we proceed with our analysis of trends in new free subscriptions. Figure 5 shows the rolling-window test result for the deterministic trend coefficient, focusing on the period before and right after the move from free to fee (January 1, 2004). Throughout 2003, new free subscriptions grew (positive trend) and showed accelerated growth in fall 2003. However, there was a decline of the deterministic trend estimate in December 2003, representing the saturating-growth regime in the work of Pauwels and Hanssens (2007). Consistent with literature (e.g., Pauwels and Hanssens 2007; Schendel, Patton, and Riggs 1976), management makes a major change (the move from free to fee) when the deterministic trend estimate is declining (i.e., when the growth prospects from the traditional advertising income model are diminishing). Further tests confirm that the move from free to fee is affected by the rolling-window trend estimate and that the effect is negative (i.e., the trend decline significantly influences the move from free to fee). Afterward, we observe a further dip in the trend estimate, reflecting the loss of (potential) free subscriptions quantified in our VARX analysis.

FIGURE 5
Rolling-Window t-Statistic of Deterministic Trend of New Free Subscriptions



Notes: The trend coefficient is significantly different from zero in each rolling window.

FIGURE 6
Rolling-Window t-Statistic of Stochastic Trend of New Free Subscriptions

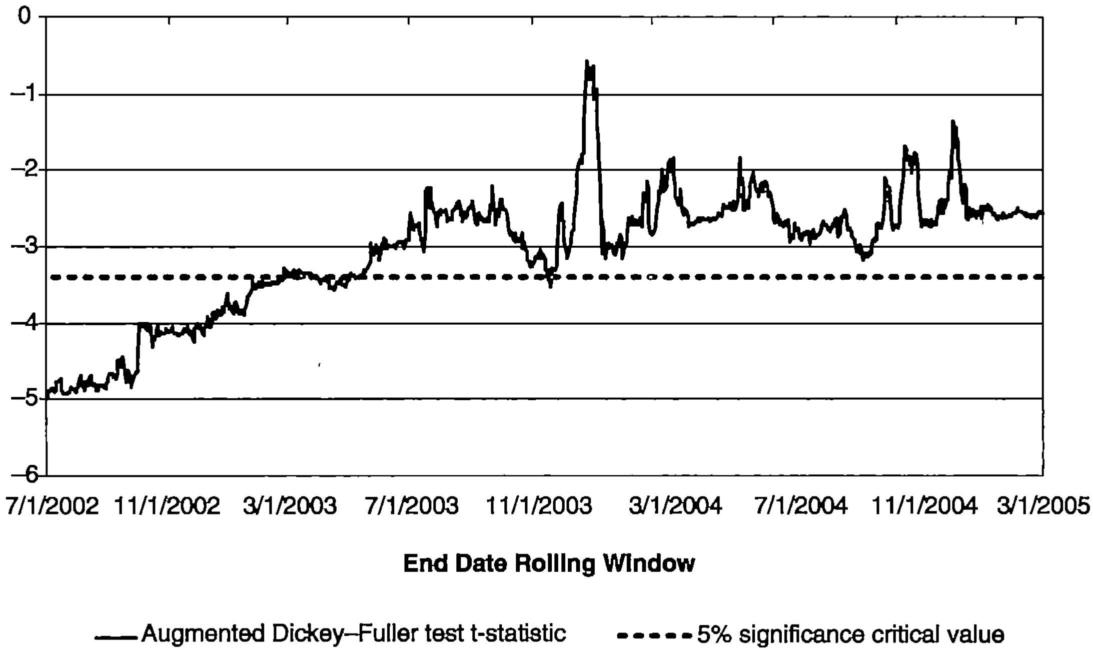


Figure 6 shows the rolling-window test results for the stochastic trend t-statistic. The stochastic trend analysis both confirms the deterministic trend assessment and highlights another aspect of the move's timing. First, the stochastic trend's t-statistic also declines in December 2003, the month before the move from free to fee. As a result, new free subscriptions are losing their classification as an evolving series and, thus, the ability to benefit permanently from a temporary subscription boost. Such change has been observed for the growth in online news consumers in Spain, marking the stabilization of consumer penetration and, thus, the need to tap other sources of growth for firms (Pauwels 2001). The move from free to fee (January 1, 2004) shows up as a return to the unit root classification because it represents a structural break in the series. Second, in contrast to previous work on growing brands (Bronnenberg, Mahajan, and Vanhonacker 2000), we do not observe a positive evolution from the start of our data set. Instead, new free subscriptions are initially classified as stationary, which changes only in spring 2003. Thus, it appears that the performance series gained "momentum" (ibid.) before management decided to move from free to fee. This influx of new free subscriptions increases the consumer value of joining the site, which enables (free and paid) users to exploit the joint expertise of the full installed base. This may be an important reason for the revenue success of the move for this particular company. It succeeded in first creating (in spring 2003) and then monetizing the positive evolution in new free subscriptions. We speculate that the paid-subscriber gains are lower and the free-subscriber losses are higher for companies that attempt to make the move before the momentum in free subscriptions has materialized. Evidently, data on such companies are needed to provide empirical evidence in this regard.

Conclusions and Avenues for Further Research

This research aims to contribute to filling an important gap in empirical research. When companies decide to move from free to fee, what is the role of marketing actions in the free- and fee-user response, and what sources of revenue loss might offset the paid subscription revenue gains? Although this trade-off is prevalent for online content, it is also likely relevant for many other marketing contexts in which users become accustomed to enjoying a product or service for free. Our analysis points specifically to the nature of the trade-offs firms face. Moving from free to fee has consequences for a firm's advertising revenues from free subscribers. The empirical results show that the move itself (plausibly through restriction of free content) slows down the growth of free users directly and reduces the effectiveness of marketing communications in generating new free users. Our content provider expected the former but not the latter, which is not readily revealed by inspection of the raw data, even after the move. In our empirical application, the revenue loss appeared minor: approximately 2% of the paid-subscriber revenue. Notably, the majority of this loss was due to the targeted e-mail blasts by the content provider aimed at convincing potential subscribers to commit to paid contracts. It appears that "content should be free" is still a powerful slogan on the Web; whereas some companies are trying to get surfers used to paying for content, other services, such as Congoo.com, are actively encouraging them to get beyond the "pay wall."

With subscription fees becoming the major revenue component for online content, the power of marketing actions to stimulate paid subscriptions is critical. Various marketing actions stimulate potential users to adopt specific pricing formats. In particular, price promotions are more effective in stimulating subscriptions for short-term contracts (in our case, monthly contracts) than for long-term contracts (in our case, yearly contracts). This finding is consistent with theoretical arguments that short-term contracts allow such users to try out the product at a minimal commitment, diagnose the extent to which the content is useful, and thus facilitate their adoption decision (Rogers and Shoemaker 1971).

In contrast to price promotions, both e-mails and search-engine referrals are more effective in generating new yearly subscriptions. The latter came as a surprise to our online provider, which suspected that the same-page display of competing content providers in search-engine results would reduce the commitment of subscribers gained though this route. However, this result is consistent with the notion that long-term commitments are riskier than short-term commitments because they involve a greater immediate economic investment. Prior research indicates that in conditions in which outcomes are risky, people look for cues that suggest that risk is minimal. One such cue is the credibility of the content provider, which may be derived through the search engine acting as a third party (Chu and Chu 1994). Likewise, the e-mails provide specific information on content, which tends to reduce perceived risk.

Finally, our moving-window trend analyses revealed that the move from free to fee occurred about a year after new free subscriptions became an evolving variable and about a month after its growth showed signs of saturation. The former is consistent with management intuition that the content site picked up momentum in signing on new users, which gave the firm the confidence to consider different ways to monetize user interest. Although such momentum may provide the opportunity to make the move from free to fee, the saturation in growth may have indicated the necessity to begin adding subscription revenues. In general, we speculate that (1) the success of the move from free to fee depends on the perceived content value of the site, which is indicated by momentum in free subscribers, and (2) management's decision to execute the move is more likely when this growth saturates. We note that saturating growth is still a relatively favorable regime in the framework of Pauwels and Hanssens (2007), who observe that packaged goods manufacturers tend to wait for dramatic changes until their performance is clearly in decline. Thus, management of a small company in a growth industry appears to act more proactively than management in established companies and mature industries (Schendel, Patton, and Riggs 1976).

This article investigated one company. Although this company is representative of the U.S. online content industry, further studies are needed to establish whether our findings generalize to other content providers and other countries. In particular, because the online provider appears to have successfully migrated to a paid-content format, it is useful to consider which factors drove these results. Consumer, competitive, and company considerations may all play a role (see Table 9).

First, the customer base growth at the time of the move from free to fee matters. Our rolling-window analysis revealed that the move occurred after the free subscription series showed a strong growth trend (i.e., momentum). Management believed that investment in search-engine optimization well before the move from free to fee was a strong driver of subscription growth, and indeed we find permanent effects of search-engine referrals. As a result, the fee subscription membership was marketed to a large and rapidly growing group of free subscribers. Other content providers may have made this switch too early (e.g., before a substantial free membership base was created), consistent with the difficulty faced by some of the earliest movers for fee-based content. For example, Microsoft entered the online magazine business with *Slate*, which unsuccessfully tried to charge subscriptions and now exists as a free site. In addition, the subscription-based nature of the industry may matter; whereas 97% of online business content is subscription based compared with single sales, this situation differs for other categories, such as online music and greeting cards. Customer reaction to the move from free to fee may well differ in transactional versus subscription-oriented markets. Finally, individual customer considerations include price sensitivity and the perceived value of free content and fee content, which are all likely to be distributed heterogeneously in the (potential) customer population.

TABLE 9
Potential Drivers on the Revenue Effects of Moving from Free to Fee

	Subscription Revenue: Paying Subscribers	Advertising Revenue: Free Subscribers
Customer growth momentum	+	+
Customer perceived value free content	– (?)	+
Customer perceived value fee content	+	– (?)
Customer price sensitivity	–	–
Competitive intensity	–	–
Competitive price level	+	?
Company content restriction extent	+	–
Company price level	–	+ (?)
Company communication effectiveness	+	+
Company free-to-fee conversion efforts	+	–
Company revenue flow management	+	+

Second, competition may be an important driver of the revenue results of moving from free to fee. Negative effects follow from a high perceived value of competitive content and its overlap with the free and fee content of the focal company, whereas a high competitive price may actually benefit the focal company. Our data provider faces a large number of competitors, some with substantial financial resources, even within the niche of marketing content. Most of these competitors follow an advertising income model, and one competitor recently announced a subscription model at a price that is an order of magnitude higher than that in our data. Several price cuts followed, presumably because of market resistance, whereas our content provider has succeeded in upgrading many consumers to higher price levels. It is unclear whether our empirical results of net revenue gains would generalize to situations with a different focal firm position (e.g., a marginal versus dominant player), substantially more versus fewer major competitors, and the choice of free versus for-fee content.

Third, the company's marketing-mix decisions may exert an important influence on the revenue outcome of moving from free to fee. Our results already point to the loss in free subscribers due to marketing actions aimed at converting free subscribers into paying ones. In addition, the small loss in new free subscriptions may be due to the content restriction decision; our content provider restricted only a portion (approximately one-seventh) of the content and maintained a steady flow of interesting free content. Furthermore, the price structure and level should matter; our data provider focuses on winning long-term (yearly) contracts and charges a yearly price close to the industry average for online business content. In contrast to the aforementioned competitor, the focal company managed to maintain prices while growing paid subscriptions. Further research is needed to confirm and operationalize the recommendation that "the key is for companies [providing Internet services] to set realistic fees and then gradually raise them" (*USA Today* 2002). Moreover, the growth in paid subscriptions is fairly well explained by several model variables, such as search-engine referrals and e-mail advertising, which are widely used. However, the extent to which these are effective depends, in large part, on how well they are executed. In addition, firms use various other techniques to move free users to paid users, such as nonprice promotional offers and advertising-based "day passes," which our content provider does not employ. These can be the subject of further research. Finally, other possible reasons for our content provider's success include its effectiveness in hiring the right people to manage the transition and to manage the resultant cash flows.

In summary, compared with the results for our data provider, we would expect lower revenue benefits for companies that (1) make the move from free to fee before momentum in free subscriptions has materialized, (2) set prices too high compared with consumer willingness to pay for their content, (3) are up against a dominant competitor with better (perceived) content and/or lower price levels, (4) charge fees for all (previously free) content, and (5) fail to ramp up marketing communication efforts and execute them effectively.

Apart from the generalizability conditions, we note several areas for further research. First, our proxy search-engine referrals should be replaced with marketing control variables that drive this outcome variable. Data limitations also prevent us from analyzing the actions of competing content providers and their effects on the focal company, a situation typical for these kinds of company data sets. Second, we do not have usage data, which would allow us to assess whether the move from free to fee also reduces the frequency with which free users visit the site and the number of pages they view. Likewise, we lack data on renewal and switch patterns for online content providers. As to our modeling approach, its focus on the aggregate level precludes an analysis of customer heterogeneity (e.g., in willingness to pay for content). Finally, our models are reduced form, and thus the long-term impact calculations are subject to the assumption that the basic data-generating process does not change. Although this is appropriate for "innovation accounting" (i.e., an account of what happened with the move from free to fee within our data sample; see Franses 2005; Van Heerde, Dekimpe, and Putsis 2005), the modeling approach is not suited for revealing structural aspects of customer and firm behavior.

A great deal of marketing research has examined how customers react to pricing when they have a reference price greater than zero. This has been valuable to our understanding of how customers react to products and services for which they expect to pay. However, it is possible that future generations may expect certain products and services, especially those delivered by the Internet, to be free, and thus they will have a reference price of zero. As content providers on the Internet have found, this is a daunting possibility. Researchers can help further illuminate how to move people from free to fee. This research is just a first step in that direction.

REFERENCES

Abraham, Magid M. and Leonard M. Lodish (1987), "Promoter: An Automated Promotion Evaluation System," *Marketing Science*, 6 (2), 101–123.

Ailawadi, Kusum L., Donald Lehmann, and Scott Neslin (2001), "Market Response to a Major Policy Change in the Marketing Mix: Learning from Procter & Gamble's Value Pricing Strategy," *Journal of Marketing*, 65 (January), 44–61.

Ariely, Dan and Kristina Shampan'er (2004), "Tradeoffs Between Costs and Benefits: Lessons from 'the Price of 0,'" working paper, Massachusetts Institute of Technology.

Bakos, Yannis and Erik Brynjolfsson (2000), "Bundling and Competition on the Internet," *Marketing Science*, 19 (1), 63–82.

Bawa, Kapil and Robert Shoemaker (2004), "The Effects of Free Sample Promotions on Incremental Brand Sales," *Marketing Science*, 23 (3), 345–63.

Bradley, Efron and Robert Tibshirani (1993), *An Introduction to the Bootstrap*. London: Chapman and Hall.

Brehm, Jack W. (1966), *A Theory of Psychological Reactance*. New York: Academic Press.

Bronnenberg, Bart J., Vijay Mahajan, and Wilfried Vanhonacker (2000), "The Emergence of Market Structure in New Repeat-Purchase Categories: The Interplay of Market Share and Retailer Distribution," *Journal of Marketing Research*, 37 (February), 16–31.

Brynjolfsson, Erik, Yu Hu, and Michael D. Smith (2003), "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science*, 49 (11), 1580–96.

Bucklin, Randolph E. and Catarina Sismeiro (2003), "A Model of Web Site Browsing Behavior Estimated on Clickstream Data," *Journal of Marketing Research*, 40 (August), 249–67.

CENTRAL LIBRARY
UNIVERSITY OF CALCUTTA
ALIPORE CAMPUS

Chu, Wujin and Woosik Chu (1994), "Signaling Quality by Selling Through a Reputable Retailer: An Example of Renting the Reputation of Another Agent," *Marketing Science*, 13 (2), 177–89.

Cooke, Alan D.J., Harish Sujan, Mita Sujan, and Barton A. Weitz (2002), "Marketing the Unfamiliar: The Role of Context and Item-Specific Information in Electronic Agent Recommendations," *Journal of Marketing Research*, 39 (November), 488–97.

Danaher, Peter J. (2002), "Optimal Pricing of New Subscription Services: Analysis of a Market Experiment," *Marketing Science*, 21 (2), 119–38.

Dekimpe, Marnik and Dominique M. Hanssens (1999), "Sustained Spending and Persistent Response: A New Look at Long-Term Marketing Profitability," *Journal of Marketing Research*, 36 (November), 397–412.

——— and ——— (2004), "Persistence Modeling for Assessing Marketing Strategy Performance," in *Cool Tools in Marketing Strategy Research*, Donald Lehmann and Christine Moorman, eds. Cambridge, MA: Marketing Science Institute.

Diehl, Kristin, Laura J. Kornish, and John G. Lynch (2003), "Smart Agents: When Lower Search Costs for Quality Information Increase Price Sensitivity," *Journal of Consumer Research*, 30 (June), 56–71.

Dyson, Esther (1995), "Intellectual Value," *Wired Magazine*, (July), (accessed December 10, 2007), [available at http://www.wired.com/wired/archive/3.07/dyson.html].

Enders, Walter (2003), *Applied Econometric Time Series*, 2d ed. New York: John Wiley & Sons.

Fitzsimons, Gavan J. and Donald R. Lehmann (2004), "Reactance to Recommendations: When Unsolicited Advice Yields Contrary Responses," *Marketing Science*, 23 (1), 1–16.

Franses, Philip-Hans (2005), "On the Use of Econometric Models for Policy Simulation in Marketing," *Journal of Marketing Research*, 42 (February), 1–14.

Gedenk, Karen and Scott A. Neslin (1999), "The Role of Retail Promotion in Determining Future Brand Loyalty. Its Effect on Purchase Event Feedback," *Journal of Retailing*, 75 (4), 433–59.

Ghysels, E., H. Lee, and J. Noh (1994), "Testing for Unit Roots in Seasonal Time Series," *Journal of Econometrics*, 62 (2), 415–42.

Granger, Clive W. and Paul Newbold (1986), *Forecasting Economic Time Series*, 2d ed. New York: Harcourt Brace Jovanovich.

Haga, Susan (2005), "Readers' Opinions," USC Annenberg Online Journalism Review, (accessed December 12, 2005), [available at http://www.ojr.org/ojr/blog/comments.cfm?id=771].

Hamilton, J. (1994), *Time Series Analysis*. Princeton, NJ: Princeton University Press.

Hanssens, Dominique M. and Ming Ouyang (2001), "Hysteresis in Market Response: When Is Marketing Spending an Investment?" working paper, Anderson School of Management, University of California, Los Angeles.

Jain, Sanjay and P.K. Kannan (2002), "Pricing of Information Products on Online Servers: Issues, Models, and Analysis," *Management Science*, 48 (September) 1123–42.

Johansen, Søren, Rocco Mosconi, and Bent Nielsen (2000), "Cointegration Analysis in the Presence of Structural Breaks in the Deterministic Trend," *Econometrics Journal*, 3 (2), 216–49.

Kwiatkowski, Dennis, Peter Phillips, Per Schmidt, and Yongcheol Shin (1992), "Testing the Null Hypothesis of Stationary Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?" *Journal of Econometrics*, 54 (1–3), 159–78.

Lambrecht, Anja and Bernd Skiera (2006), "Paying Too Much and Being Happy About It: Existence, Causes, and Consequences of Tariff-Choice Biases," *Journal of Marketing Research*, 43 (May), 212–23.

Laurent, Gilles and Jean-Noel Kapferer (1985), "Measuring Consumer Involvement Profiles," *Journal of Marketing Research*, 22 (February), 41–53.

Leeflang, Peter S.H., Dick R. Wittink, Michel Wedel, and Philippe A. Naert (2000), *Building Models for Marketing Decisions*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Lewis, Regina (1995), "Relation Between Newspaper Subscription Price and Circulation, 1971-1992," *Journal of Media Economics*, 8 (1), 25–41.

Lütkepohl, Helmut (1993), *Introduction to Multiple Time Series Analysis*. New York: Springer-Verlag.

Maddala, G.S. and In-Moo Kim (1998), *Unit Roots, Cointegration, and Structural Change*. Cambridge, UK: Cambridge University Press.

Neslin, Scott A. and Robert W. Shoemaker (1983), "Using a Natural Experiment to Estimate Price Elasticity: The 1974 Sugar Shortage and the Ready-to-Eat Cereal Market," *Journal of Marketing*, 47 (Winter), 44–57.

Online Publishers Association (2003), *Online Paid Content: U.S. Market Spending Report*, (accessed January 18, 2006), [available at http://www.online-publishers.org/pdf/opa_paid_content_report_mar03.pdf].

——— (2005), *Online Paid Content: U.S. Market Spending Report*, (accessed January 18, 2006), [available at http://www.online-publishers.org/pdf/opa_paid_content_report_fullyear05.pdf].

Pauwels, Koen (2001), "The Drivers of Online Readership: A Decomposition Approach," *Social Science Research Network* Working Paper No. 01-08.

——— (2004), "How Dynamic Consumer Response, Competitor Response, Company Support and Company Inertia Shape Long-Term Marketing Effectiveness," *Marketing Science*, 23 (4), 596–610.

——— (2007), "How Retailer and Competitor Decisions Drive the Long-Term Effectiveness of Manufacturer Promotions for Fast-Moving Consumer Goods," *Journal of Retailing*, 83 (3), 297–308.

——— and Enrique Dans (2001), "Internet Marketing the News: Leveraging Brand Equity from Marketplace to Marketspace," *Journal of Brand Management*, 8 (4–5), 303–314.

——— and Dominique M. Hanssens (2007), "Performance Regimes and Marketing Policy Shifts," *Marketing Science*, 26 (3), 1–19.

———, ———, and S. Siddarth (2002), "The Long-Term Effects of Price Promotions on Category Incidence, Brand Choice, and Purchase Quantity," *Journal of Marketing Research*, 34 (November), 421–39.

———, Shuba Srinivasan, Jorge Silva-Risso, and Dominique M. Hanssens (2004), "New Products, Sales Promotions, and Firm Value: The Case of the Automobile Industry," *Journal of Marketing*, 68 (October), 142–56.

Perron, Pierre (1989), "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57 (6), 1361–1401.

——— (1990), "Testing for a Unit Root in a Time Series with a Changing Mean," *Journal of Business & Economic Statistics*, 8 (2), 153–62.

Pesaran, M. Hashem, Richard G. Pierse, and Kevin C. Lee (1993), "Persistence, Cointegration, and Aggregation: A Disaggregated Analysis of Output Fluctuations in the U.S. Economy," *Journal of Econometrics*, 56 (1–2), 57–88.

——— and Yongcheol Shin (1998), "Generalised Impulse Response Analysis in Linear Multivariate Models," *Economics Letters*, 58, 17–29.

Peterson, Robert A., Sridhar Balasubramanian, and Bart J Bronnenberg (1997), "Exploring the Implications of the Internet for Consumer Marketing," *Journal of the Academy of Marketing Science*, 25 (4), 329–46.

Rogers, Everett M. and Floyd F. Shoemaker (1971), *Communication of Innovations: A Cross-Cultural Approach*. New York: The Free Press.

Saba, Jennifer (2005), "Dispelling the Myth of Readership Decline," *Editor & Publisher* (online edition), (November 28), (accessed January 3, 2006), [available at http://www.editorandpublisher.com/eandp/news/article_display.jsp?vnu_content_id=1001571463].

Schendel, Dan E., Richard G. Patton, and James Riggs (1976), "Corporate Turnaround Strategies. A Study of Profit Decline and Recovery," *Journal of General Management*, 3 (Spring), 3–11.

Scott, Carol A. (1976), "The Effects of Trial and Incentives on Repeat Purchase Behavior," *Journal of Marketing Research*, 13 (August), 263–69.

Sims, Christopher A. and Tao Zha (1999), "Error Bands for Impulse Responses," *Econometrica*, 67 (5), 1113–56.

Swanson, Norman R. and Halbert White (1997), "Forecasting Economic Time Series Using Flexible Versus Fixed Specification and Linear Versus Nonlinear Econometric Models," *International Journal of Forecasting*, 13, 439–61.

Telang, Rahul, Peter Boatwright, and Tridas Mukhopadhyay (2004), "A Mixture Model for Internet Search-Engine Visits," *Journal of Marketing Research*, 41 (May), 206–214.

USA Today (2002), "Internet Firms Fish for Revenue from Fees," (June 11), (accessed December 10, 2007), [available at http://www.usatoday.com/money/covers/2002-06-11-tech-sell.htm].

Van Heerde, Harald J., Marnik G. Dekimpe, and William P. Putsis Jr. (2005), "Marketing Models and the Lucas Critique," *Journal of Marketing Research*, 42 (February), 15–21.

———, Sachin Gupta, and Dick Wittink (2003), "Is 75% of the Sales Promotion Bump Due to Brand Switching? No, Only 33% Is," *Journal of Marketing Research*, 40 (November), 481–91.

Varian, Hall (1995), "The Information Economy," *Scientific American*, (September), 200–202.

Wertenbroch, Karl (1998), "Consumption Self-Control by Rationing Purchase Quantity of Virtue and Vice," *Management Science*, 17 (4), 317–37.

Zeithaml, Valarie A. (1988), "Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence," *Journal of Marketing*, 52 (July), 2–22.

Zellner, Arnold and Henri Theil (1962), "Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations," *Econometrica*, 30 (1), 54–78.

Ajay Kalra & David A. Soberman

The Curse of Competitiveness: How Advice from Experienced Colleagues and Training Can Hurt Marketing Profitability

The literature suggests that mangers are overly competitive and overemphasize competitors' payoffs in making decisions. The authors demonstrate how two seemingly innocuous but common decision inputs that mid- and entry-level managers use—(1) advice from colleagues and (2) training—could propagate this bias. The authors conduct an experiment in which one generation of participants plays a noncooperative game and then passes on advice on how to play the same game to the second generation. The authors find that the second generation is more likely to accept decisions that negatively impact the competitor's profits, even when those decisions result in lower payoffs. The results suggest that the advice of the first generation causes the second generation to be more sensitive to relative performance. In a second experiment, the authors show that training films that create negative feelings toward competitors can have adverse effects on managers' decisions. The training increases the willingness of managers to sacrifice own profits to reduce competitor profitability. These findings have important implications for managers in competitive markets. This research suggests that standard training materials that emphasize beating competitors should be carefully assessed for their unintentional but potentially damaging effects. In particular, training designed to evoke affective responses toward competitors must be used judiciously. Firms that give junior managers more autonomy and responsibility should also be cognizant that the advice these managers get and the information they share may propagate a competitive bias. To mitigate this problem, an option is to devote resources to better understand the content and nature of advice that managers share with one another.

Keywords: corporate training, planning process, knowledge transfer, dynamic games, noncooperative games, competition

The effectiveness of marketing spending often falls short of expectations (Santucci 2004). Much of the lack of effectiveness has been attributed to poor decision making caused by "too many things to think about," "not having the right information," and/or an inability to recognize what Grove (1996) calls "strategic inflection points" (see also Wind and Crook 2005). To date, the focus has been primarily on decisions and errors that organizational leaders, company presidents, and top executives make. However, the decisions of regional brand managers, trade marketers, and key account managers within a company receive little attention from academics and the popular press. Such employees may not be high profile, but their combined impact on the effectiveness of marketing is high. Typically, mid- and entry-level managers' major decisions include the allocation of promotional budgets, regional pricing, and setting sales force incentive schemes.

Several explanations have been put forth in the academic literature to account for poor decision making by managers, such as insufficient information to make decisions carefully (e.g., Christen, Boulding, and Staelin 2004) and bounded rationality (e.g., Simon and Kaplan 1989). Behavioral decision theory also provides a further explanation for biases that belie managerial rational decision making (e.g., Boulding, Morgan, and Staelin 1997; Chakravarti, Mitchell, and Staelin 1981). The specific bias we investigate is managers' tendency to derive utility not only from their own profits but also from negative payoffs to competitors.

Managers' tendency to select strategies that are overly competitive is widely documented and has its origins in early studies of the "prisoner's-dilemma" game (Axelrod 1984). In identifying strategies that lead to the highest outcome in two-player repeated prisoner's-dilemma tournaments, Axelrod's (1984) work demonstrates that tit-for-tat yields the highest outcomes. Fader and Hauser (1988) extend this work to tournaments with three players and continuous decisions. These studies show that managers do not naturally gravitate to strategies that maximize outcomes. Griffith and Rust (1997) extend these ideas by demonstrating that managers place a high value on performance relative to competitors (and not absolute profits), even when they are explicitly instructed to maximize own profits and are compensated on the basis of their own profit performance. In a survey of existing evidence, Armstrong and Green (2007) emphasize the danger of competitor-oriented objectives. The recent behavior of John Mackey, chief executive officer (CEO) of Whole Foods, who, over several

---

Ajay Kalra is an associate professor, Tepper School of Business, Carnegie Mellon University (e-mail: kalra@andrew.cmu.edu). David A. Soberman is an associate professor, INSEAD (e-mail: david.soberman@insead.edu). The authors thank Dave Reibstein and Ting Zhu for their input on this project. The order of authorship is alphabetical.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

years, posted numerous messages on Yahoo Finance praising Whole Foods and disparaging competitors, such as archrival Wild Oats, further underscores the prevalence of excessively competitive behavior (*The Wall Street Journal* 2007).

Because the overemphasis of competitive payoffs in decision making has potentially detrimental effects for the firms and shareholders, identifying the genesis of this bias and factors that engender it is of theoretical and substantive interest. Several factors, such as education before employment, cultural backgrounds, and personality, can potentially account for why managers are prone to this competitive bias. However, Armstrong and Collopy (1996) find that the bias exists in 43 experimental replications across several samples tested in different countries. Because the bias appears to be fairly widespread, we examine the inputs managers use to make decisions as possible causes of this predilection.

Prior literature has essentially viewed these inputs as either data or decision rules (e.g., Chakravarti, Mitchell, and Staelin 1981). Because decision-making inputs do not explain why managers accentuate competitors' profits, we identify other inputs that are salient and have a significant effect in the decision-making context of the junior manager. The objective of this article is to demonstrate how two seemingly innocuous but common decision inputs—advice from colleagues and training—could propagate this bias.

To understand how advice from colleagues and training constitute a possible genesis of this bias, we examine the context in which entry-level managers make decisions. A first characteristic of this environment is the time pressure that managers face to learn (and understand) a business and to make decisions. A study conducted by Accenture indicates that the average time it takes a company to create and launch a marketing campaign is 2.5 months. Moreover, 58% of the executives surveyed stated that they have attempted to shorten this time (*Business Wire* 2001).

A second characteristic that affects decision making is the level and quality of information that managers have to make decisions. Surveys indicate that marketing managers do not have access to the information they need to make effective decisions. According to the Accenture study, obtaining accurate and up-to-date data to make decisions is a critical need for 68% of managers (*Business Wire* 2001). Procter & Gamble's chief marketing officer, Jim Stengel, stresses this issue by commenting that "[m]arketing is a $450 billion industry, and we are making decisions with less data and discipline than we apply to $100,000 decisions in other aspects of our business" (VanBoskirk 2005, p. 2). Even if research is available, for several reasons, managers perceive the information as lacking insight relevant to decision making (e.g., Hellebusch 2004; Mittal 2004). The problems associated with conducting and using marketing research are also well documented (e.g., Andreasen 1985).

A third characteristic that affects marketing decisions is that the relationship between marketing decisions and marketing outcomes is not well established (e.g., Rust et al. 2004). The key implication of these characteristics is that mangers may not have adequate information or conduct sufficient analysis to develop effective marketing plans. As a result, managers use other information and approaches to make decisions.

One source of information that junior managers use is advice or guidance from their peers or colleagues in similar positions (e.g., Schotter and Sopher 2007). Unlike CEOs, midlevel managers have easy access to many colleagues with the requisite experience who are often in adjacent offices. The benefit of advice from colleagues is that these people have made similar decisions in the past. Seeking advice or information from others is common across many domains.[1] Indeed, both in and outside the workplace, it is natural for people to receive guidance from others, even when they do not actively seek it.

A second source of guidance for decision making is cues from the environment. A cue that is most relevant for junior managers is the "rules of action" learned from the training new and midlevel employees receive early in their tenure with a company. This cue is salient for mid- and entry-level managers but not for senior executives; the majority of training within firms is directed toward junior managers. Training often reflects the beliefs and culture of a company and provides concrete information as to how to make decisions.

A natural question is to ask how these sources of guidance affect decision making. In particular, is there a link between the overcompetitiveness of managerial decisions and the sources of guidance mid- and entry-level managers use to make decisions? Both advice from colleagues and training seem useful for transmitting marketing knowledge and expertise to managers who are new. Our objective is to understand these sources of guidance and highlight their unintentional but potentially pernicious effects. Using an experimental approach, we examine the decision making of participants who use these two sources of guidance.

Conceptual Development

A natural question arises as to why managers and firms promote a culture that focuses on inflicting damage on competitors. There are several possible explanations. First, there are situations, such as a zero-sum game, in which decreasing the competitors' profits is entirely rational. In addition, when competitors interact with each other in a repetitive sense, inflicting harm on "the competitors" may be profit reducing in the short run but profit maximizing in the long run. This can be the case when the short-term profit-reducing action is part of a tit-for-tat strategy that allows players to sustain a cooperative outcome in a dynamic game (Kreps et al. 1982). Alternatively, inflicting harm on the competitor may maximize profit in the long run if the competitor is ultimately eliminated (Parks, Pharr, and Lokeman 1994). Nevertheless, it seems reasonable to argue that most market situations are not really zero-sum games; in general, increased competition results in lower profits for firms and more benefits for customers (Boynton, Blake, and Uhl 1983; Brodie, Bonfrer, and Cutler 1996; Leeflang and Wittink 1996). In addition, firms are rarely successful in eliminating their key competitors.

---

[1] Such behavior is frequently encouraged through explicit statements of corporate values. For example, Hewlett-Packard, through its "HP Way" statement of corporate values (and, subsequently, through its "Rules of the Garage"), strongly encourages employees to adopt an open-door policy to discuss business conduct and to share "tools and ideas" (see the "HP E-Services.Solutions" Case EC-16 of the Graduate School of Business, Stanford University, May 2000).

Second, there is a body of research demonstrating that coordination is an important factor in a collective setting to achieve desired outcomes (Grandori and Soda 2000; Schmitt, Swope, and Walker 2000). Because demonizing the competition provides managers with a common message and a common set of values, it may be useful to improve coordination within a firm. To the best of our knowledge, there is no empirical study that directly measures the impact of culture on effectiveness. Nevertheless, Hofstede (1980), Schein (1985), and Weick (1987) propose and find evidence that culture coordinates collective action taking.

Third, it may be easier to describe the set of actions that harm a competitor than to describe actions that increase profit. As Armstrong and Collopy (1996) note, competitor-oriented objectives may be useful because a competitor's performance can serve as a benchmark. In this context, instilling a culture that emphasizes the importance of hurting the competition will assist managers in identifying the relevant information and the most effective action. Brown and Starkey (1994) suggest that a key effect of organizational culture is to specify what information is important and thus assist in focusing a manager's attention.

Fourth, it may be easier to convince managers to take action that inflicts harm on a competitor than action that improves profit. Research shows that people are distributed along a continuum with highly competitive (or individualist) personalities at one end and highly cooperative personalities at the other (Chatman and Barsade 1995; Kelley and Stahelski 1970). People close to the highly competitive end of the continuum derive significant benefit from competition and beating an opponent. There is evidence that salespeople who score high on the competitiveness trait set higher goals and realize superior performance (Brown, Cron, and Slocum 1998). Because the competitive trait positively affects performance, it is plausible that actions that enhance these tendencies also lead to superior results. This perspective of personality provides an explanation for the popularity of boxing and World Wrestling Entertainment; a major component of these sports is the demise of artificial villains (Baldwin and Flaherty 1999). If people drawn to careers in marketing and sales are strong on competitive traits, it may be easier to motivate them by focusing on actions that equate competitors to a common foe. There is evidence that the fit between personality type and organizational culture is a key determinant of behavior (Chatman and Barsade 1995).

Finally, research on contests has demonstrated that merely participating in a contest provides utility for the players (e.g., O'Keefe, Viscusi, and Zeckhauser 1984). Because sales contests can be effective to motivate employees by having them compete with each other (e.g., Kalra and Shi 2001), managers may frame the competitive nature of business as a contest. As the English essayist Charles Lamb (1823) succinctly put, "Man is a gaming animal. He must always be trying to get the better in something or other."

Both advice from colleagues and training are important and useful sources of information that junior managers use to make marketing decisions. We believe that these sources of information also lead managers to make decisions that place a high weight on harming competitors. As a result, the ability of marketing spending to generate returns in terms of profit is adversely affected. Next, we discuss these two sources of guidance.

Advice

A source of guidance for new managers is knowledge obtained from more experienced colleagues (e.g., Schotter and Sopher 2003). As Day and Nedungadi (1994) note, managers' decisions are subject to psychological processes learned through experience and shared as industry conventional wisdom. Managers also use informal tools (proverbs, procedures, and perspectives) to convey wisdom that exists within the organization (Klayman, Larrick, and Heath 2000). This is particularly true of managers who pass on advice to colleagues or subordinates who inherit their responsibilities over time. The evidence from experimental economics shows that "word-of-mouth" social learning between successive generations of managers is an important force in the creation of social conventions (Schotter and Sopher 2007).

Our purpose is to understand how managers are affected by advice from other managers who have experience in making marketing decisions. As we noted previously, managers across many backgrounds are prone to use competitors' negative payoffs in their decision making (Armstrong and Collopy 1996; Griffith and Rust 1997). Unlike perceptual and psychological biases, this bias seems to be propagated within the firm. In particular, the process of advice giving suggests that a first generation of managers may have a detrimental effect on a second generation of managers who are placed in an identical decision situation. If the behavioral tendencies of the first generation are amplified in the second generation, a rationale for why this bias spreads and endures over time is provided. Therefore, we predict the following:

> H_1: Advice from a previous generation of experienced players is more likely to lead to decisions that create negative payoffs for competitors.
>
> H_2: Advice from a previous generation of experienced players is more likely to lead to emphasis on relative payoffs than on absolute payoffs.

Training

Another path of knowledge transfer is the training and socialization processes that junior managers undergo. Some of the major objectives of training are skill development, enhancement of organizational commitment, and team or coalition building (e.g., Tannenbaum and Yukl 1992). The content of training and motivational programs that firms use often includes material that evokes affective responses toward competitors. For example, one firm gives a "shoot-

’em-up” computer game to employees, in which the objective is to destroy targets that are identified by competitors’ names and logos. Another training film simulates a war between the firm and its major competitor. In addition, senior management reinforces an enemy-like portrayal of competitors. The computer chip maker Cyrix is known to have placed Intel’s tombstone in its lobby atrium, and the CEO of Cabletron allegedly concluded his pep talks to sales recruits by plunging a knife into a beach ball inscribed with a competitor’s name (Fierman and Barlyn 1995). Similarly, the CEO of PepBoys used videotapes of burning and burying baseball caps inscribed with competitor logos. We argue that such training processes evoke changes in both the cognitive and the affective components of attitudes managers have toward their competitors.

Bruch and Ghoshal (2003, p. 46) argue that emotions play a key role in how firms operate. They categorize organizations in terms of energy zones, one of which is the aggression zone. They propose that competitors in the aggression zone operate on “internal tensions founded on strong negative emotions.” There is accumulating evidence that emotions are important in an organizational setting (e.g., Diefendorff and Richard 2003; Fisher and Ashkanasy 2000), though the evidence on whether they have a positive or negative impact is mixed (e.g., Morris and Feldman 1997).

In consumer behavior, the affective state, particularly feelings, has been demonstrated to influence the processing of information, evaluations, and judgments (e.g., Garg, Inman, and Mittal 2005; Lee and Sternthal 1999). Our analysis of training films indicates that many of them portray competitors as enemies. These stimuli may evoke negative affect toward the competitors, which in turn affects the managers’ decision processes. Specifically, this may lead to a decision process in which increased importance is placed on creating negative payoffs for competitors.

H_3: Training that increases negative affect toward competitors is more likely to lead to decisions that create negative payoffs for competitors, even at a cost to own profitability

Experiment 1

The purpose of Experiment 1 is to investigate the effect of advice from experienced colleagues on new managers who are in a situation of making marketing decisions. In situations in which learning processes are important in explaining how actors behave over time, experiments with successive generations are useful to understand and predict the behavior of markets (Lucas 1986). Our goal is to examine the decision-making performance of successive generations of managers who face similar scenarios in a competitive, noncooperative context. We adopt a modified version of Merlo and Schotter’s (1999, 2003) experimental approach.

Participants were asked to play a noncooperative game with four alternative strategies. The payoffs were based on simultaneous decisions made by the player and the opponent (the payoff matrix appears in Figure 1). The standard prisoner’s-dilemma experiment entails two choices for the players such that the dominant strategy leads to a pareto-inferior outcome for both players. To understand situations in which the complexity of coordination is higher, games with more than two choices are used experimentally (Cooper et al. 1990; Holt 1985).

On completing the experimental procedure, each participant was asked to provide advice on how to play the game. They were told that this advice would be passed to the “next manager” of the firm (i.e., the next generation of participants). The second-generation participants played the same game, the only difference being that they were provided with additional information in the form of written advice from the previous generation.[2] The participants were compensated on the basis of their performance in the simulation; however, those in the first generation were also compensated on the basis of the performance of the next manager. We compare the decisions of the two generations of players.

To simulate many real markets, the opponent was designed to be cooperative some of the time, but not always. The objective of the experiment was to observe how the first-generation managers responded to their experience. Given the design of the incentive structure, a profit-maximizing “rational” approach would suggest that the first generation would recommend strategies that maximize profits. The first generation shared in the profits of the second generation, so it was in the best interest of the first gen-

FIGURE 1
The Payoff Matrix

		Competitor's Decision			
		Low	**Medium**	**High**	**Max**
Your Decision	**Low**	20, 20	10, 20	0, 20	–40, 10
	Medium	20, 10	30, 30	10, 20	–20, 10
	High	20, 0	20, 10	20, 20	–10, 0
	Max	10, –40	10, –20	0, –10	0, 0

Notes: The number on the left in each cell represents the participant's payoff, and the number on the right represents the competitor's payoff.

---

[2] This approach is similar to that which Schotter and Sopher (2001, 2003) employ to study the role of advice on behavior in a context of nonoverlapping generations of decision makers.

eration to provide advice to improve the performance of the second generation.

Procedure

Sixty-four students participated in the experiment for extra course credit and monetary payments based on performance in the experiment. Each session consisted of six to eight participants. At the beginning of the task, the participants were instructed to take the role of a marketing manager making advertising decisions. They were told that their firm was competing against another firm managed by another participant. They were also told that they would not be able to identify their competitor. Participants were given the task of deciding between alternative sizes of advertising budgets for each quarter. The advertising budget was described as "low," "medium," "high," or "maximum" (max), indicating higher amounts of dollars spent. Participants were informed that the competitor would simultaneously make the same decision. The payoffs were based on the combination of advertising budgets chosen. The participants were then provided the payoff information of Figure 1. They were told that their competitors' payoffs were identical. The procedure and lab were structured so that the participants could not communicate with one another. To retain the guise that participants were playing against others, the game was programmed to move to the next decision simultaneously for all participants after each decision.

After each round, the screen displayed information on the participant's decision, the competitor's decision, the results, and cumulative payoffs for both players. The payoff matrix was displayed before each round. The participants played two practice rounds, after which they played eight rounds. To avoid endgame effects, they were not told how many rounds they would play. The game took approximately 30 minutes. After all decisions were made, the participants were asked to assume that they had been moved to another position but needed to give advice to their successor. To ensure that the advice task was taken seriously, each participant was told that his or her payoff (in real money) would be partially based on the performance of the successor as well as his or her own performance. The participants were paid $.05 for each unit of profit earned. Participants in the first generation were paid their earnings as well as the amount their successors made (the payoff of the second generation was based only on own performance).

There were 35 participants in the first generation and 29 in the second generation. Because of technical problems, we did not obtain complete data for 2 participants in the first generation and 1 participant in the second generation. Therefore, we do not use the data for these participants in the analysis. The only difference in the procedure between the first- and the second-generation participants was that the second generation received a hard copy of advice from a manager in the first generation, whereas the first generation received no advice.

The opponent was computerized and preprogrammed to make quasi-random choices as a function of the choices made by the participant. From the perspective of the participant, the opponent's choices appeared to be made simultaneously. The distribution of choices was made according to the probability distributions shown in Figure 2. We used a preprogrammed, computerized opponent for three reasons. First, because each participant faced an identical opponent, the comparison of behavior across participants is facilitated. Second, the random decision rule the computerized opponent employed ensured that participants could not quickly determine the decision process of the opponent. Third, the probability distribution reflects an opponent that is not oriented toward cooperation (the opponent deviates from the cooperative outcome of medium–medium).

Game Structure

The game was presented to the participants as a game of simultaneous decision making in which the number of possible actions was fixed at four. As we discussed previously, a game with more than two strategies is useful when the researcher wants to add complexity beyond the typical 2 × 2 decision task used to analyze the repeated prisoner's dilemma (Iyengar and Schotter 2008; Merlo and Schotter 1999, 2003).

Given the payoff matrix of Figure 1, the optimal action for the participant in terms of profit maximization is a function of the action the opponent takes. At the beginning of the game, we assumed that the participant was unsure of how the opponent would play. As a result, we considered three possible bases a participant might use to make decisions. Each basis is a function of the beliefs the participant might have about how the opponent makes decisions.

1. The participant believes that the opponent is rational.

From Figure 1, it is straightforward that the pareto-optimal outcome in the game is (medium–medium). This is the strategy combination in which each manager earns 30 units. If both managers play medium, neither has an incentive to deviate because this is the highest payoff that can be earned.

FIGURE 2
The Programmed Responses of the Computerized Opponent

Participant's Choice	Randomized Competitor's Response
Low	33% Low 33% Medium 34% High 0% Max
Medium	0% Low 33% Medium 33% High 34% Max
High	0% Low 33% Medium 33% High 34% Max
Max	34% Low 0% Medium 33% High 33% Max

Note that all symmetric strategies are Nash equilibria; however, in a repeated game with rational players, (medium–medium) would be expected to be the long-term expected outcome. Thus, there is a strong rationale to play "medium" in anticipation that the opponent will cooperate.

2. The participant believes that the opponent is random.

Because the opponent plays stochastically, a participant may infer that the opponent plays in a completely random way. In this case, the expected payoff for the participant is based on the opponent choosing each advertising level with a probability of 25%. We can calculate the expected payoff of strategy i as follows:

$$\Pi_i = \sum_{j = L, M, H, X} \frac{1}{4}\pi(s_j|i).$$

The expected payoffs for low, medium, high, and max are –2.5, 10, 12.5, and 5, respectively. This implies that the optimal response to a competitor playing randomly is "high."

3. The participant believes that the opponent behaves according to the quasi-random decision rule in Figure 2.

It is unlikely, but conceivable, that after a few rounds, the participant correctly infers the quasi-random decision rule of the opponent. This would first entail the participant realizing that the opponent's decision is not made simultaneously to his or hers. The participant would then need to realize that the opponent's decision depends on his or her decision. Given the number of rounds played in the experiment, it is unlikely that the participant would make these realizations. Nevertheless, this basis constitutes an important benchmark against which participants' actual decisions can be compared. Using the conditional probabilities of Figure 2, we calculate the expected outcome for each participant's strategy choices using the following formula:

$$\Pi_i = \sum_{j = L, M, H, X} p(s_j|i)\,\pi(s_j|i).$$

The expected payoffs for low, medium, high, and max are 10, 6.66, 10, and 3.33, respectively. This implies that the optimal response to a quasi-random competitor who makes decisions according to Figure 2 is either low or high. The worst response a participant can make given these beliefs is max.

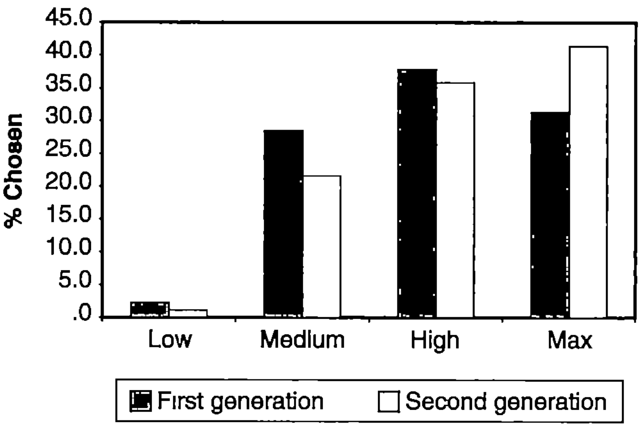
This discussion demonstrates that the optimal action for the participant depends entirely on his or her beliefs about the opponent's next decision. The game is constructed so that learning creates a natural migration in the beliefs of the participant. Over the course of several decisions, a rational participant might migrate from the first basis to the second basis and from the second basis to the third basis. Depending on a participant's belief, an argument can be made for choosing low, medium, or high. A critical characteristic of the game is that only the belief that the opponent will play max with 100% probability provides a profit-maximizing justification for a participant to choose max. This belief is unlikely to be sustained, given that we observed a string of more than four successive decisions of max (by the opponent) only once in 640 rounds of the experiment.[3]

The practice of including a dominated strategic choice with unique characteristics (e.g., max) in an experimental setting is not new. For example, Abdalla and colleagues (1989) include an additional strategy in each player's strategy set in the context of trying to understand the tendency of players to act cooperatively (i.e., altruistically), even when such action can be costly. The additional strategy had the property of being payoff dominated, but it also led to the highest joint payoff if both players chose it. In contrast, our objective is to evaluate the tendency of players to act aggressively, even when such action leads to inferior outcomes. The max strategy is a strategy that has the property of being payoff dominated, but it also leads to the lowest performance for the competitor, independent of (the competitor's) strategy.

Results

We present the distribution of decisions for the first and second generations in Figure 3. The striking difference between the two generations is that the frequency of the max decision is much higher in the second generation. Recall that there is no profit-based rationale for choosing max, given the payoffs to the participants. The figures also demonstrate that the participants focus their decisions on the top three levels of advertising. Although a decision of low has the potential to generate higher profits, it appears that participants were reluctant to choose it because of its potential to deliver minimum payoffs (–40). Overall, the results suggest that the second-generation managers were affected by the advice of the first-generation managers. Their advice appears to stimulate profit-reducing choices on the part of the second generation.

FIGURE 3
Distribution of Choices by Generation



---

[3] Given the preprogrammed strategy of the competitor, the maximum probability of a string of n max's is $.33^n$. In the 640 decisions that competitors made in the experiment, only 1 string of 5 max decisions was observed, and only 2 strings of 4 were observed. All other strings of consecutive max decisions were 3 or less.

H_1 predicts that advice will lead to decisions that emphasize more negative payoffs for competitors. To test the hypothesis, we first examine whether the overall frequency of choices made is significantly different across the two generations. The results indicate that the frequency of choices is significantly different ($\chi_1^2 = 8.02$, $p < .05$). Next, we consider two alternative measures of the participants' decisions. We define max or high as "aggressive" decisions and medium or low as "nonaggressive" decisions.

With an aggressive decision, the participant reduces the likelihood that the opponent will choose an advertising level that is higher (that beats the participant). More important, an aggressive action reduces the expected payoffs for the competitor. The distribution of decisions according to this metric appears in Table 1. The results indicate that the second generation made significantly more aggressive decisions than the first generation ($\chi_1^2 = 4.75$, $p < .05$). Thus, we reject the hypothesis that the number of aggressive choices made in the two generations is identical.

We also consider another measure of the nature of decisions that reflects participants' tendency to act aggressively even when it almost certainly results in lower profits. We define an advertising decision of max as being "very aggressive." With a very aggressive decision, a participant eliminates the likelihood that the opponent will choose an advertising level that is higher. In addition, a decision of max guarantees that the opponent's profits cannot exceed zero.

Again, the results (Table 1) indicate that the second generation made a significantly higher proportion of very aggressive decisions than the first generation ($\chi_1^2 = 5.62$, $p < .05$). Overall, the results support H_1; that is, the second generation, who received advice, behaved differently from the first generation. In particular, second-generation participants were more likely to make decisions that negatively affected the competitor's profits, even when those decisions resulted in lower payoffs for themselves. Next, we analyze the data to obtain insights into the decision processes of the first- and second-generation decision makers.

Decision Process

H_2 predicts that the second generation of managers will be more attentive than the first generation to the competitors' profit performance. In particular, strong competitor performance should provide an incentive for the participant to take more aggressive or very aggressive actions in the second generation.

To test H_2, we examine the proneness to make an aggressive or a very aggressive decision as a function of two independent measures of the outcome of the previous decision: own payoffs and the competitor's payoffs (Payoff_{t-1} represents the participant's payoff lagged by one period, and Competitor payoff_{t-1} is the competitor's payoff lagged by one period). We examine the main effects and interaction effects with Gen, a dummy variable that reflects the generation (1 for the first generation and 0 for the second generation). The significance of the interaction parameters provides an indication of the degree to which the second-generation managers are more sensitive to their own profit performance and that of the competitor. In summary, the second generation should be more likely to consider competitors' payoffs and make more aggressive choices if the competitors' payoffs were higher in the previous period. Thus, we estimate the probability of making an aggressive (very aggressive) decision with a binary logit model; we show the attractiveness of the decision in Equation 1:

$$(1)\quad \begin{aligned} V_{it} = \exp(&\beta_0 + \beta_1 \text{Payoff}_{t-1} + \beta_2 \text{Competitor payoff}_{t-1} \\ &+ \beta_3 \text{Gen} + \beta_4 \text{Payoff}_{t-1} \times \text{Gen} \\ &+ \beta_5 \text{Competitor payoff}_{t-1} \times \text{Gen} + \varepsilon_{it}), \end{aligned}$$

where V_{it} is the attractiveness of an aggressive (very aggressive) choice by participant i at time t and ε_{it} is a random error term that follows the extreme value distribution.

Equations 2 and 3 are the estimated models, and $p(\text{agg}_{it})$ and $p(\text{veryagg}_{it})$ are the probabilities that participant i makes an aggressive or very aggressive choice in period t, respectively:

$$(2)\quad p(\text{agg}_{it}) = \frac{1}{1 + V_{it}}, \text{ and}$$

$$(3)\quad p(\text{veryagg}_{it}) = \frac{1}{1 + V_{it}}.$$

We summarize the results in Table 2, Panels A and B. The results for the aggressive decision estimation reveal one significant main effect. Participants' decisions were positively affected by the size of their own payoffs in the previ-

TABLE 1
Distribution of Choices Across Generations

		First Generation	Second Generation	Total
Decision	Nonaggressive	95	62	157
		(30.7%)	(22.7%)	26.7%
	Aggressive	214	211	425
		(69.3%)	(77.3%)	73.0%
	Total	309	273	582
				(100%)
Decision	Not very aggressive	212	160	372
		(68.6%)	(58.6%)	63.9%
	Very aggressive	97	113	210
		(31.4%)	(41.4%)	36.1%
	Total	309	273	582
				(100%)

TABLE 2
Experiment 1: Using Prior Payoffs as Explanatory Variables

A: Estimates for Aggressive Decisions

Parameter	Estimate	SE	χ^2
Intercept β_0	−2.3073	.3124	54.56**
$\text{Payoff}_{t-1}\ \beta_1$	.0767	.0201	14.49**
Competitor $\text{payoff}_{t-1}\ \beta_2$	.0171	.0120	2.03
Gen β_3	.2520	.4017	.39
$\text{Payoff}_{t-1} \times$ Gen β_4	−.0563	.0254	4.92*
Competitor $\text{payoff}_{t-1} \times$ Gen β_5	.0530	.0212	6.24*

B: Estimates for Very Aggressive Decisions

Parameter	Estimate	SE	χ^2
Intercept β_0	−.0146	.1570	.01
$\text{Payoff}_{t-1}\ \beta_1$	.0464	.0116	15.91**
Competitor $\text{payoff}_{t-1}\ \beta_2$	.0041	.0076	.29
Gen β_3	.2924	.2150	1.85
Payoff × Gen β_4	−.0065	.0160	.16
Competitor payoff × Gen β_5	.0283	.0107	6.93**

*$p < .05$.
**$p < .01$.

ous period ($p < .001$), indicating that they were more likely to choose an aggressive decision if they realized strong performance in the previous period. This result affirms that participants were strategic in their decision making in terms of balancing own profits and minimizing competitor profits; they were more likely to engage in lowering the competitor's payoffs after accumulating their own profits. Note from Figure 1 that aggressive decisions could also imply a risk-averse posture in the trade-off between lowering losses and attempting large gains. The main effect of generation in the model for aggressive decisions is nonsignificant.

The most germane findings in the aggressive estimation are the significant interaction effects between generation and own/competitor payoffs in the previous period (for both estimated parameters, $p < .05$). First, the interaction effect with own payoffs in the previous period is negative. This implies that the decisions of participants in the second generation were less sensitive to their own payoffs; the sum of β_2 and β_4 represents the combined effect of own payoffs on making an aggressive decision for the second generation. Although the combined effect remains positive, it is lower in magnitude for the second generation. Second, and more important, participants in the second generation were more likely to make aggressive decisions when the competitor realized good performance in the previous period. Overall, the aggressive estimation suggests that the second generation was less focused on its own performance and more focused on lowering the competitor's profit performance than the first generation.

The story in the very aggressive estimation is somewhat different but corroborates the dynamic suggested in the aggressive estimation. Similar to the results of the aggressive decision, participants were significantly more likely to choose a very aggressive decision when their own payoffs were higher in the previous period ($p < .001$). Most important, the interaction term between generation and competitor payoffs is both positive and significant ($p < .05$). This implies that participants in the second generation were significantly more likely to choose a strategy that guaranteed poor competitor profits when the competitor recently performed well. A distinguishing feature of a very aggressive choice is that it guarantees poor profits for the competitor. In contrast, a choice of high (one of the two possible choices that define an aggressive decision) allows for the possibility of good profits for the competitor (the competitor can earn 20 if he or she also chooses high). In summary, the very aggressive estimation shows that the second generation is significantly more interested in limiting the competitor's performance.

Together, the results of the two models support H_2; that is, the second generation is less affected by own prior performance but more affected by the desire to drive the profits of the competitor downward. We now specifically test whether the second generation focuses to a greater extent on performance relative to the competitor than on absolute performance. To evaluate this, we estimate a model based on the difference in profits between the participant and the competitor in the previous period: Relative $\text{payoff}_{t-1} = \text{Payoff}_{t-1} - \text{Competitor payoff}_{t-1}$. Again, we estimate the probability of making an aggressive (very aggressive) decision with a binary logit model; the attractiveness of choosing an aggressive (very aggressive) decision is as follows:

$$(4)\quad V_{it} = \exp(\beta_0 + \beta_1\delta_{t-1} + \beta_2\text{Gen} + \beta_3\delta_{t-1} \times \text{Gen} + \varepsilon_{it}).$$

The results for $p(\text{agg}_{it})$ and $p(\text{veryagg}_{it})$ appear in Table 3, Panels A and B. Both models reveal that the difference in profits has an effect on the probability that participants in the second generation will choose aggressive ($p < .10$) and very aggressive ($p < .05$) decisions, respectively. When the difference in profits is lower in the previous period, partici-

TABLE 3
Experiment 1: Using Relative Prior Payoff as an Explanatory Variable

A: Estimates for Aggressive Decisions

Parameter	Estimate	SE	χ^2
Intercept β_0	–1.4131	.1781	62.95***
Relative payoff$_{t-1}$ β_1	–.0025	.0086	.08
Gen β_2	.3136	.2304	1.85
Relative payoff$_{t-1}$ × Gen β_3	–.0219	.0124	3.11*

B: Estimates for Very Aggressive Decisions

Parameter	Estimate	SE	χ^2
Intercept β_0	.2008	.1424	1.99
Relative payoff$_{t-1}$ β_1	.0086	.0069	1.52
Gen β_2	.4280	.1951	4.82**
Relative payoff$_{t-1}$ × Gen β_3	–.0221	.0096	5.36**

*$p < .10$.
**$p < .05$.
***$p < .01$.

pants in the second generation are more likely to choose aggressive and significantly more likely to make more very aggressive decisions.

To explore the impact of advice on performance, we compare the average profits of the first and second generation. The average profit earned by the first generation was 7.72 (14.08) units per round, whereas the second generation earned only 6.23 (13.91) units per round. This represents a drop of more than 19%, which is marginally significant ($p <$ .10). The performance of the computerized competitor across generations is also notable. In the first generation, the competitor earned 4.10 (20.27) units per round, and in the second generation, it earned 1.39 (20.01) units per round. This represents a drop of more than 65%, which is significant ($p < .05$). These findings show that the second generation reduced the performance of the competitor effectively but did so at a significant cost to own performance. Thus, the advice seems to cause the second generation to be significantly more sensitive to relative performance and, in particular, to focus on driving the competitor's profits lower.[4]

Nature of Advice

To examine the nature of the advice and its impact on the second generation in greater depth, we conducted a protocol analysis. We developed a coding scheme after examining the first set of advice. The advice was categorized into six main categories: cooperative, hard-line (aggressive and hostile), derogatory, adaptive, and reactive. First, we coded advice that recommended cooperating with the competitor to maximize profits mutually as "cooperative advice." Second, we coded advice that recommended aggressive behavior (e.g., using only max or max/high decisions) as "hard-line advice." Third, we coded the subset of hard-line advice that advocated aggressive behavior without explicitly stating that it should only be in response to competitors' decisions as "aggressive advice." We coded another subset of hard-line advice that recommended specifically to outsmart/minimize competitors' payoffs as "hostile advice." Fourth, we coded advice that made derogatory or negative comments about the competitor as "derogatory comments." Fifth, we coded advice as "adaptive advice" if it encouraged the successor to respond to the, competitors' decisions. Sixth, we coded advice as "reactive" if it counseled that the competitor decisions would react to decisions made by the successor.

Two judges who were blind to the objective of the study put each sentence of the advice into one of these six categories. Then, the advice was coded as 1 if it contained at least one sentence that fit any of the categories and 0 if it did not. The interjudge reliability was 88%, and we resolved all disagreements.

The results show that only 14.29% of the participants suggested cooperative behavior, whereas 60.71% recommended hard-line behavior. Furthermore, 39.29% of the participants advised their successors to be aggressive, regardless of competitor decisions. Only 10.71% of the participants made specific derogatory or negative comments about their competitors, though 21.43% suggested that the competitor should be outsmarted. Finally, 46.43% of the advice recommended that the game be played adaptively on the basis of the competitors' decisions, and 21.43% provided guidance that competitors would react to the successor's decisions.

We then examined whether specific advice received by the second-generation managers affected their decisions. The distribution of aggressive and very aggressive decisions made by those who received aggressive advice and those who did not appears in Table 4. The results show that the nature of advice significantly affects decision making. Par-

---

[4] As we noted previously, maximizing relative performance is identical to maximizing absolute performance in a zero-sum game. However, in most marketing situations, maximizing relative performance invariably implies diminished absolute performance.

TABLE 4
Distribution of Choices in Second Generation

		Nonaggressive Advice	Aggressive Advice	Total
Decision	Nonaggressive	44 (28.21%)	18 (15.38%)	62 22.71%
	Aggressive	112 (71.79%)	99 (84.62%)	211 77.29%
	Total	156	117	273 (100%)
Decision	Not very aggressive	104 (66.67%)	56 (47.86%)	160 58.61%
	Very aggressive	52 (33.33%)	61 (52.14%)	113 41.39%
	Total	156	117	273 (100%)

ticipants in the second generation who were advised to be aggressive made significantly more aggressive decisions ($\chi_1^2 = 6.26$, $p < .05$) than those who did not get such advice, and they also made significantly more very aggressive decisions ($\chi_1^2 = 9.74$, $p < .01$).

The proportion of hard-line advice is partially reflective of the first-generation managers' negative attitudes toward the competitors. We speculate that the inherent competitive bias is compounded by the competitors' actions. The negative outcomes for the first generation are caused by the interplay between the inherent competitive bias and the deviations from the equilibrium medium–medium decision by the computerized opponent. The saliency bias suggests that negative events are noticed and overweighted more (e.g., Ito et al. 1998). The negativity of the first generation results in advice directed at creating a negative attitude toward the competitors. In addition, some of the advice explicitly advocated taking preemptive aggressive actions. This is consistent with previous research that demonstrates how emotions can affect judgments (Johnson and Tversky 1983; Wright and Bower 1992).

Experiment 2

The purpose of Experiment 2 is to investigate the effects of training on new managers who are in a pressurized situation to make marketing decisions. We exposed participants to stimuli that are similar to the materials used to train managers at sales meetings and annual conferences. The stimuli were designed to manipulate the affective component of attitudes toward a competitor. We then asked participants to make a series of marketing decisions.

The participants were 57 managers who were in an executive program. They were asked to participate in a "marketing decision" study after completing the marketing module of the program. Their average work experience was 7.5 years, their average age was 32 years, and 30% were women.

Participants were asked to assume that they had been hired by Molson brewery to manage the Coors Light brand in Canada. They were told that they would make marketing decisions after undergoing a training program conducted by the firm. As part of the training program, they were given background material about the market and viewed a short training film. In the background information, participants were provided information about the market structure, firm performance, and excerpts from fictitious interviews given by the CEOs of two firms. They were told that there were four major players in the market and that the firms competed through advertising, merchandising, and major event sponsorships rather than price. They were also informed that Molson's major competitor was Labatt. The performance of Molson in the last year was described as poor, and it was noted that the profits had been below expectations. After reading the materials and seeing the film, participants made a series of marketing decisions for Molson. After participants made their decisions, we performed a manipulation check on their feelings. Finally, we collected demographic information.

Manipulations

The factor we manipulated was attitude toward competitors (either neutral or negative). The experimental design was a between-subjects, single-factor design with two levels. Feelings toward the key competitor were manipulated with two elements: the content of the background information and the training film. First, part of the background information consisted of excerpts from a *Wall Street Journal* article in which both the chairman of Molson and the chairman of Labatt (Molson's chief competitor) were interviewed. In the neutral condition, the chairman of Molson stated that he was "cautious but optimistic" about next year. The chairman of Labatt commented that "we have worked extremely hard this year and were tremendously helped by our dedicated employees. He hoped that the results next year would also be very good." In the negative-affect condition, the chairman of Molson stopped short of allegations but criticized Labatt for "not playing by the rules." The chairman of Labatt stated that "we followed a very aggressive strategy this year and have done better than our competitors. They (our competitors) are not very good and behind the times—we plan to beat them this year by a wide margin. I promise that they'll have a lot to complain about next year."

The second component of the manipulation was in the content of the training film. In the neutral condition, participants viewed a film in which the sports marketing plan of the Coors brand was discussed. The film highlighted plans to promote the brand by organizing innovative baseball, skiing, and golf events that would potentially increase market penetration by creating brand excitement. The role of each event was discussed. In the negative-affect condition, the training film showed employees of Molson dressed in army gear attacking Labatt. In three short vignettes, the Molson team blow up the competitor's products, destroy a Labatt supply truck, and finally dynamite the Labatt headquarters. Both films were edited versions of training films actually used by Molson to motivate its sales teams.

We tested the two training videos with a similar sample of 88 participants to ensure that they were equivalent with respect to (1) the overall reaction to the training video and (2) the objective of aiding in training-related motivation (e.g., Mathieu, Tannebaum, and Salas 1992; Noe 1986). In addition, because a major objective in marketing training films is to enhance employees' affective judgments toward their employer, we also measured attitudes toward the firm. Participants' reactions to training assessed both the affective reaction—"What is your overall reaction to the training video?" (anchored by "strongly negative" and "strongly positive")—and the perceived relevance of the video to the job—"How valuable was the content of the training video to the marketing job at Coors/Molson?" (Cronbach's $\alpha = .75$). We measured attitudes toward the firm by asking participants to assess three statements on seven-point Likert scales: "The video makes me feel good about being a marketing manager at Molson/Coors," "The video makes me feel that Molson/Coors is a good place to work," and "The characters in the video convey a friendly image of what it is like to work at Molson/Coors" (Cronbach's $\alpha = .81$). We measured participants' training-related motivation by asking them to rate two statement on seven-point Likert scales: "The types of actions I am expected to take to make my business successful are made clearer by a video like this," and "This video will result in me performing higher quality work at Molson/Coors" (Cronbach's $\alpha = .81$).

The results indicate that the overall reactions to both the training videos were not significantly different ($\bar{X}_{\text{Neutral}} = 4.36$ versus $\bar{X}_{\text{Negative Affect}} = 4.30$; $F_{1,87} = .05$, n.s.). Furthermore, both the training videos were perceived as equivalent to training-related motivation ($\bar{X}_{\text{Neutral}} = 4.92$ versus $X_{\text{Negative Affect}} = 4.74$; $F_{1,87} = .51$, $p < .48$). Finally, as we expected, the negative-affect training video increased positive attitudes toward the employer firm compared with the neutral video ($\bar{X}_{\text{Neutral}} = 4.46$ versus $\bar{X}_{\text{Negative Affect}} = 5.02$; $F_{1,87} = 4.04$, $p < .05$). Overall, the results suggest that the two training videos were similarly relevant to the marketing job and were perceived as equally motivating.

Dependent Variables

Participants were given three marketing scenarios and were asked to make a decision in each. The three scenarios required the participants to select an advertising campaign, a merchandising program, and an event sponsorship. In each scenario, they were given four alternatives and were asked to rank order their preferred options. For each scenario, participants were provided with the expected outcomes for their own firm and for their three main competitors: Carling, Labatt, and Pacific Western. The key outcomes are for the focal firm, Molson, and the major competitor, Labatt. The stimuli appear in Table 5. In all scenarios, the outcomes were constructed such that one alternative was "maximize own and main competitor outcome," in which both Molson and Labatt made the maximum possible profits. In all scenarios, this option is Alternative A. For example, in Scenario 1, which is the advertising expenditures decision, Alternative A provides both Molson and Labatt \$70 million—the maximum profit both firms can earn among the four alternatives. Similarly, in Scenario 2, Alternative A provides both Molson and Labatt the maximum impact scores of 29.6 and 26.5, respectively. The second alternative is classified as "maximize own outcome/hurt competitor." In this alternative, which is Alternative B in all scenarios, the outcome for the own firm, Molson, is as high as the first alternative, but the outcome for the main competitor, Labatt, is lower than the first alternative. To illustrate, in Scenario 2 (merchandizing display), Alternative A provides the maximum impact for Molson (29.6) and for Labatt (26.5). In Alternative B, Molson continues to get the maximum impact of 29.6, but the impact for Labatt is reduced from 26.5 to 14.5. The third alternative "sacrifice profit/attack competitor," which is Alternative C in all scenarios, is constructed such that the profits for the own firm, Molson, decrease slightly from the maximums of Alternatives A and B, but the profits for the competitor firm decrease sharply. For example, in Scenario 3, the impact score for Molson decreases from the maximum of 8 to 7.5 if Alternative C is selected. However, the impact scores for Labatt drop to 1 versus 7 or 6 if Alternative A or B is selected, respectively. The fourth alternative, Alternative D, is a "dummy choice" in which all other options dominate. We included the dummy choice to check whether the participants were carefully evaluating the alternatives. To avoid order effects, we rotated the position of various options across participants' marketing decisions.

Manipulation Check

After participants completed the marketing decisions for the three scenarios, we measured feelings toward the main competitor using six five-point scales that asked participants to state how they felt about Labatt. Given the objective of the study, we measured only negative feelings. We selected the items from the negative feelings scale that Edell and Burke (1987) developed. The six items were critical, angry, disgusted, defiant, offended, and suspicious. We asked participants to indicate the extent to which they felt each of the items on a scale ranging from 1 ("not at all strongly") to 5 ("very strongly"). We formed an index of feelings by averaging the scores for these items (Cronbach's $\alpha = .80$). The results indicate that the manipulation was successful. Participants in the negative-affect condition had significantly greater negative feelings toward the competitor ($X_{\text{Negative Affect}} = 3.01$) than those in the neutral condition ($X_{\text{Neutral}} = 2.45$; $F_{1,56} = 3.71$, $p < .05$). Although significantly different, the feelings in the negative-affect condition were not high on the absolute scale. This is not surprising, because negative feelings toward an entity are usually

TABLE 5
Experiment 2 Stimuli

A: Scenario 1: Advertising Expenditures

Beer Companies	Campaign A	Campaign B	Campaign C	Campaign D
Your profit (Molson)	$70	$70	$68	$60
Carling	$50	$30	$65	$55
Labatt	$70	$50	$10	$60
Pacific Western	$30	$50	$65	$45

B: Scenario 2: Merchandising (Displays in Retailers) Impact Scores

Beer Companies	Canada Day Displays (A)	Victoria Day (B)	Sounds of Summer Displays (C)	Christmas/New Year Displays (D)
Your impact score (Molson)	29.6	29.6	29.1	27.9
Carling	22.0	20.0	28.7	22.0
Labatt	26.5	26.1	14.5	26.1
Pacific Western	23.0	21.9	27.7	22.0

C: Scenario 3: Major Event Sponsorships

Beer Companies	Rodeo Sponsorship (A)	Snowmobiling Sponsorship (B)	Fishing Sponsorship (C)	Horse-Racing Sponsorship (D)
Your impact score (Molson)	8	8	7.5	6
Carling	5	6	4	5
Labatt	7	6	1	6
Pacific Western	2	3	4	3

evoked by experience rather than by short-term role playing.

Results

H_3 predicts that managers exposed to training material that creates negative affect are more likely to sacrifice own profits to create negative payoffs for competitors. To test H_3, we conducted a logit analysis using the condition and the demographics of years of experience and sex. The percentage of participants selecting each choice in the advertising decisions appears in Table 6. We created two dependent measures. The first dependent measure is the proportion of times the sacrifice profit/attack competitor choice was selected as the first or second preference. This is useful as a dependent measure to quantify the fraction of non-profit-maximizing behavior by each participant. Recall that in each scenario, there are two choices available that equally maximize profit for the decision maker. The implication of the sacrifice profit/attack competitor choice being selected first or second is that the participants preferred the sacrifice profit option to one or both of the profit-maximizing decisions. Therefore, this measure is an indicator of the level of utility the managers obtain from lower profits accruing to the competitor. The results of the logit analysis appear in Table 7, Panel A. We find strong support for H_3. The analysis shows that manipulating the training materials had a significant impact on the decision. In all three scenarios, there was a significant difference in the proportion of participants who elected to sacrifice their own profits to reduce their competitor's profitability (for Scenarios 1 and 2, $p < .05$, and for Scenario 3, $p < .10$).

Not surprisingly, the most popular first choice was the maximize own profits/hurt competitor option. This choice allowed the participant to maximize profit and inflict harm

TABLE 6
Results of Experiment 2

	Scenario 1		Scenario 2		Scenario 3	
	Neutral	Negative Affect	Neutral	Negative Affect	Neutral	Negative Affect
Maximize own and main competitor outcome	20.00%	.03%	23.33%	11.11%	26.67%	29.63%
Maximize own outcome/hurt competitor	66.67%	60.00%	60.00%	40.74%	33.33%	22.22%
Sacrifice profit/attack competitor	13.33%	26.67%	16.66%	48.15%	36.67%	44.44%
Dummy	.00%	.00%	.00%	.00%	3.33%	3.70%

Notes: Numbers indicate proportion of participants' first preference.

TABLE 7
Dependent Variable

A: First Choice: Sacrifice Profits/Attack Competitor

	Scenario 1		Scenario 2		Scenario 3	
	Estimate	F	Estimate	F	Estimate	F
Intercept	–2.42	2.76	–2.45	2.98	.09	
	(1.46)		(1.39)		(1.28)	
Sex[a]	.43	.44	.52	.68	–.58	.88
	(.65)		(.63)		(.60)	
Experience[b]	.06	.32	.16	2.07	.03	.10
	(.11)		(.11)		(.10)	
Condition[c]	1.58	6.80***	.98	3.99**	.72	1.63*
	(.60)				(.56)	

B: First or Second Choice: Sacrifice Profits/Attack Competitor

	Scenario 1		Scenario 2		Scenario 3	
	Estimate	F	Estimate	F	Estimate	F
Intercept	–2.58	1.51	–2.40	2.98*	.09	.01
	(2.09)		(1.40)		(1.28)	
Sex	–.65	.56	.52	.68	–.58	.93
	(.86)		(.63)		(.60)	
Experience	.05	.11	.16	2.07	.03	.10
	(.17)		(.11)		(.10)	
Condition	1.88	2.83**	1.17	3.98**	.72	1.64*
	(1.12)		(.58)		(.56)	

*$p < .10$.
**$p <$ 05.
***$p < .01$.
[a] Sex is coded as 1 for males and 0 for females.
[b] Experience is in number of years.
[c] Condition is coded as 1 for negative affect and 0 for neutral.

simultaneously on the competitor. Nevertheless, in none of the scenarios was this option selected first by more than 67% of the participants. As a result, a second measure of the relative preference for sacrificing profit to inflict harm on the competitor is the frequency that sacrifice profit/attack competitor was selected first.

Using sacrifice profit/attack competitor as a first choice as the measure of how willing participants were to sacrifice profit to harm the competitor, we present the results from the logit analysis in Table 7, Panel B. Again, the results support our hypothesis. In both Scenarios 1 ($p < .05$) and 2 ($p < .05$), the manipulation led participants to place increased weight on the competitor's profits in making their decisions. In Scenario 3, the manipulation did not create significant results in terms of whether the option to sacrifice profit was selected as the first choice.

Discussion

Overall, the results of Experiment 2 support the notion that training can have adverse effects on the quality of mid- and entry-level managers' decisions. Using actual training films, the experiment demonstrates that training can evoke negative feelings toward competitors. The negative feelings increase managers' willingness to sacrifice own profits to reduce competitor profitability.

The experiment has several limitations worth noting. First, despite the use of real training films, it is difficult to achieve a replication of how training affects the decisions of managers over the long run. The manipulation of feelings created in an experiment are unlikely to be strongly held or to have an impact on decision making compared with feelings created through training that occurs repetitively over several years. Second, the stimuli were designed to evoke only negative-affective reactions. We did not examine cognitions or nonnegative arousal, such as excitement from rivalry.[5] Third, in contrast to Experiment 1, there were no monetary payoffs provided to participants. The use of managers participating in an executive program as participants poses several potential problems. For example, manipulation biases could be introduced by an instructor who precedes the experiment. We minimized the potential for this bias because the marketing module was taught after the experiment. In addition, sample biases are likely to occur if participants are not representative of the managerial population (Gordon, Slade, and Schmitt 1987). We acknowledge this limitation but believe that it poses a low threat to external validity because the sample consisted of managers drawn from 11 countries across different industries. In addition, Ashton and Kramer (1980) find similarity between students and nonstudents in studies of decision making.

---

[5] We thank an anonymous reviewer for this point.

Conclusion

The literature has primarily considered nonoptimal decision making the result of a lack of information or the decision rules that managers use. These factors are important across many decision situations. Operational decisions by midlevel managers, such as promotional allocation and regional pricing, are made under considerable informational and time constraints. We add to the literature by examining two additional inputs these managers rely on in their decision making: advice from colleagues and employee training. We show how both sources of guidance lead managers to place emphasis on relative (and not absolute) performance. In particular, we demonstrate how advice and corporate training exacerbate a problem that is well documented; namely, managers place excessive importance on the ability of strategies to generate negative outcomes for competitors. Firms need to condition mid- and entry-level managers to compete effectively. The problem is that the actions of the firm condition mid- and entry-level managers to compete excessively. Our study also underscores how the decision environment of junior managers is important for understanding why the tendency to focus on competitor payoffs is ingrained in firms.

In theory, advice from experienced colleagues to help make decisions under pressure appears both reasonable and judicious. In Experiment 1, we evaluate the effect of this process and find support that advice from colleagues contributes to the persistence and amplification of the competitor payoff bias (Armstrong and Collopy 1996; Griffith and Rust 1997). A protocol analysis of the advice reveals a negative attitude toward the competitors. Our intuition is that a general saliency bias, in which negative events are overweighted, causes this attitude. Prior literature has suggested that negative events receive more attention and that people often exaggerate the frequencies of negative events. As a result, the advice counsels an aggressive approach, largely justified by the negative emotion the first generation develops toward the competitor. In addition, in Experiment 1 (and likely in practice), deviations from cooperative equilibrium may cause a negative affect that also provokes negative feedback.

We believe that this dynamic may be amplified by organizational inertia. Some midlevel managers remain in their positions for a long time. Given their experience, these managers are often sought by colleagues for advice. It is possible that experienced midlevel managers contribute most to the effects discussed herein.

The second source of information we consider is training. Junior managers are exposed to significant training and cultural inculcation (particularly after joining) to facilitate integration into the company. Frequently, training places high importance on the value of beating competitors. This orientation for training is the logical response of companies to conventional wisdom about how salespeople should be trained. For example, a study of more than 500,000 salespeople conducted by Gallup concludes that a key characteristic of great salespeople is that they are people who "want to beat colleagues or the competition" (Brewer 1994, pp. 83–84).

Experiment 2 demonstrates that some elements of training programs overemphasize the importance of reducing the performance of competitors by creating negative affect toward competitors. Although engineering employee emotions can have several beneficial aspects, the results of our study show that firms should be cognizant of the potential drawbacks of using such content.

Our research suggests that the competitive focus of training materials needs to be reassessed. Managers join firms with a range of mind-sets (from cooperative to competitive) that may be inherent or driven by experience. Firm-specific experience, including word-of-mouth advice and training, can accentuate or attenuate these tendencies. Further research should examine the differential effectiveness of such learning or unlearning processes. A potential avenue is to examine whether training that focuses on evoking cognitive rather than affective reactions is more effective.

The design of the payoff structure in the advice experiment was motivated by the standard S-shaped advertising response function. Thus, we assume that there is an optimal level of expenditures for both firms after which increasing advertising leads to decreasing returns. Further research should explore alternative payoff structures in which there are no pareto-optimal outcomes. We assumed that the payoff structure is consistent with the participants' prior beliefs. Because the participants in Experiment 1 were students, prior beliefs were not a major concern. In further research with managers, payoff structures could be constructed that match their domain of experience. Such procedures can be used to ensure that managers do not need to unlearn their prior beliefs. It would also be useful to investigate the moderating effect of specific real-world experience and prior beliefs about payoff structures on the nature of actions taken and the advice.[6]

The insights from our study are most relevant to the context of mid- and entry-level marketing managers. As such, they should be of particular concern to companies with large marketing departments. In addition, there is a growing trend among traditionally centralized firms, such as Sara Lee and Coca-Cola, to decentralize and give local managers more autonomy and responsibility (*The Economist* 2000). This implies that junior managers make more marketing decisions than in the past. Given that useful information is also shared by mid- and entry-level managers, it is unclear how this process can be improved. Conversely, this article demonstrates that the general orientation of many training materials may cause decision making that is not in the best interest of the firm.

---

[6] We thank an anonymous reviewer for this point.

REFERENCES

Abdalla, A.R., R. Cooper, D. Dejong, R. Forsythe, and T. Ross (1989), "Forward Induction in Coordination and Battle of Sexes Games: Some Experimental Results," working paper, Carleton Industrial Organization Research Unit, Department of Economics, Carleton University, Ottawa.

Andreasen, Alan (1985), "Backward Market Research," *Harvard Business Review*, 63 (May–June), 179–82.

Argyle, Michael (1991), *Cooperation: The Basis of Sociability*. London: Routledge Press.

Armstrong, J. Scott and Fred Collopy (1996), "Competitor Orientation: Effects of Objectives and Information on Managerial Decisions and Profitability," *Journal of Marketing Research*, 23 (May), 188–99.

——— and Kersten C. Green (2007), "Competitor-Oriented Objectives: The Myth of Market Share," *International Journal of Business*, 12 (1), 115–34.

Ashton, R.H. and S.S. Kramer (1980), "Students as Surrogates in Behavioral Accounting Research: Some Evidence," *Journal of Accounting Research*, 18 (1), 1–16.

Axelrod, Robert (1984), *The Evolution of Cooperation*. New York: Basic Books.

Baldwin, Kristen and Mike Flaherty (1999), "35 Million Wrestling Fans Can't Be Wrong (or, How We Learned to Stop Worrying and Love the Pile Driver)," *Entertainment Weekly*, (April 16), 14–15.

Boulding, William F., Ruskin Morgan, and Richard Staelin (1997), "Pulling the Plug to Stop the New Product Drain," *Journal of Marketing Research*, 34 (February), 164–76.

Boynton, Robert D., B.F. Blake, and J.N. Uhl (1983), "Retail Price Reporting Effects in Local Food Markets," *American Journal of Agricultural Economics*, 65 (1), 20–29.

Brewer, Geoffrey (1994), "Mind Reading: What Drives Top Salespeople to Greatness," *Sales and Marketing Management*, 146 (May), 82–92.

Brodie, R.J., A. Bonfrer, and J. Cutler (1996), "Do Managers Overreact to Each Other's Promotional Activity? Further Empirical Evidence," *International Journal of Research in Marketing*, 3 (4), 379–87.

Brown, Andrew D. and Ken Starkey (1994), "The Effect of Organizational Culture on Communication and Information," *Journal of Management Studies*, 31 (6), 807–829.

Brown, Stephen P., William L. Cron, and John W. Slocum (1998), "Effects of Trait Competitiveness and Perceived Intraorganizational Competition on Salesperson Goal Setting and Performance," *Journal of Marketing*, 62 (October), 88–98.

Bruch, Heike and Sumantra Ghoshal (2003), "Unleashing Organizational Energy," *Sloan Management Review*, 45 (1), 45–51.

Business Wire (2001), "Sixty-Eight Percent of Executives Unable to Measure Marketing Campaign Return on Investment, Says Accenture Study," (November 7), 1.

Chakravarti, Dipankar, Andrew A. Mitchell, and Richard Staelin (1981), "Judgment-Based Marketing Decision Models: Problems and Possible Solutions," *Journal of Marketing*, 45 (Fall), 13–23.

Chatman, Jennifer A. and Sigal G. Barsade (1995), "Personality, Organizational Culture, and Cooperation: Evidence from a Business Simulation," *Adminstrative Science Quarterly*, 40 (3), 423–43.

Christen, Markus, William Boulding, and Richard Staelin (2004), "Optimal Market Intelligence Strategy When Attention Is Scarce," working paper, INSEAD.

Cooper, R., D. Dejong, R. Forsythe, and T. Ross (1990), "Selection Criteria in Coodination Games: Some Experimental Results," *American Economic Review*, 80 (1), 218–33.

Day, George S. and Prakesh Nedungadi (1994), "Managerial Representations of Competitive Advantage," *Journal of Marketing*, 58 (April), 31–44.

Diefendorff, J.M. and E.M. Richard (2003), "Antecedents and Consequences of Emotional Display Rule Perceptions," *Journal of Applied Psychology*, 88 (2), 284–94.

The Economist (2000), "Fashion Victim," (January 20), 74.

Edell, Julie A. and Marian Burke (1987), "The Power of Feelings in Understanding Advertising Effects," *Journal of Consumer Research*, 14 (3), 421–33.

Fader, Peter S. and John R. Hauser (1988), "Implicit Coalitions in a Generalized Prisoner's Dilemma," *Journal of Conflict Resolution*, 32 (3), 553–82.

Fierman, Jaclyn and Suzanne Barlyn (1995), "When Genteel Rivals Become Mortal Enemies," *Fortune*, 131 (9), 90–100.

Fisher, C.D. and N.M. Ashkanasy (2000), "The Emerging Role of Emotions in Work Life: An Introduction," *Journal of Organizational Behavior*, 21 (2), 123–29.

Garg, Nikita, J. Jeffery Inman, and Vikas Mittal (2005), "Incidental and Task-Related Affect: A Re-Inquiry and Extension of the Influence of Affect on Choice," *Journal of Consumer Research*, 32 (June), 154–59.

Gordon, Michael E., L. Allen Slade, and Neal Schmitt (1987), "Student Guinea Pigs: Porcine Predictors and Particularistic Phenomena," *Academy of Management Review*, 12 (1), 160–63.

Grandori, A. and G. Sod (2000), *Organization and Economic Behavior*, Chap. 5. Abingdon, UK: Taylor & Francis, 135–55.

Griffith, David E. and Roland T. Rust (1997), "The Price of Competitiveness in Competitive Pricing," *Journal of the Academy of Marketing Science*, 25 (2), 109–116.

Grove, Andrew S. (1996), *Only the Paranoid Survive*. New York: Bantam Doubleday Dell.

Hellebusch, Stephen (2004), "Cost Grows When Needed Research Overlooked," *Marketing News*, (September 15), 25.

Hofstede, Geert (1980), *Culture's Consequences: International Differences in Work Related Values*. Beverly Hills, CA: Sage Publications.

Holt, Charles A. (1985), "An Experimental Test of the Consistent-Conjectures Hypothesis," *American Economic Review*, 75 (3), 314–25.

Ito, Tiffany A., Jeff T. Larsen, N. Kyle Smith, and John T. Cacioppo (1998), "Negative Information Weighs More Heavily on the Brain: The Negativity Bias in Evaluative Categorizations," *Journal of Personality and Social Psychology*, 75 (4), 887–900.

Iyengar, Raghuram and Andrew Schotter (2008), "Learning Under Supervision: An Experimental Study," *Experimental Economics*, forthcoming.

Johnson, E.J. and A. Tversky (1983), "Affect, Generalization and the Perception of Risk," *Journal of Personality and Social Psychology*, 45 (1), 20–31.

Kalra, A. and M. Shi (2001), "Designing Optimal Sales Contests: A Theoretical Perspective," *Marketing Science*, 20 (2), 170–93.

Kelley, Harold H. and Anthony J. Stahelski (1970), "Social Interaction Basis of Cooperators' and Competitors' Beliefs About Others," *Journal of Personality and Social Psychology*, 16 (1), 66–91.

Klayman, J., R. Larrick, and C. Heath (2000), "Organizational Repairs," *Across the Board*, 37 (February), 26–31.

Kreps, David M., Paul Milgrom, John Roberts, and Robert Wilson (1982), "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 27 (2), 253–79.

Lamb, Charles (1823), *Essays of Elia*. Philadelphia: H.C. Carey & I. Lea.

Lee, Angela Y. and Brian Sternthal (1999), "The Effects of Positive Mood on Memory," *Journal of Consumer Research*, 26 (September), 115–27.

Leeflang, Peter S.H. and D.R. Wittink (1992), "Diagnosing Competitive Reaction Using (Aggregated) Scanner Data," *International Journal of Research in Marketing*, 9 (1), 39–57.

Lucas, Robert E. (1986), "Adaptive Behavior and Economic Theory," in *Rational Choice: The Contrast Between Economics and Psychology*, R. Hogarth and M. Reder, eds. Chicago: University of Chicago Press, 217–42.

Mathieu, J.E., S.I. Tannenbaum, and E. Salas (1992), "Influences of Individual and Situational Characteristics on Training Effectiveness Measures," *Academy of Management Journal*, 35 (4), 828–47.

Merlo, A. and A. Schotter (1999), "A Surprise-Quiz View of Learning in Economic Experiments," *Games and Economic Behavior*, 28 (1), 25–54.

——— and ——— (2003), "Learning by Not Doing; An Experimental Investigation of Observational Learning," *Games and Economic Behavior*, 42 (1), 116–36.

Mittal, Vikas (2004), "Research and the Bottom Line," *Marketing Research*, 16 (3), 36–40.

Morris, J.A. and D.C. Feldman (1997), "Managing Emotions in the Workplace," *Journal of Managerial Issues*, 9 (3), 257–74.

Noe, R.A. (1986), "Trainee Attributes and Attitudes: Neglected Influences on Training Effectiveness," *Academy of Management Review*, 11 (4), 736–49.

O'Keefe, M., W.K. Viscusi, and R.J. Zeckhauser (1984), "Economic Contests: Comparative Reward Schemes," *Journal of Labor Economics*, 2 (1), 27–56.

Parks, Bill, S.W. Pharr, and B.D. Lokeman (1994), "A Marketer's Guide to Causewitz: Lessons for Winning Market Share," *Business Horizons*, 37 (July–August), 68–73.

Rust, Roland, Tim Ambler, Gregory Carpenter, V. Kumar, and Rajendra Srivastava (2004), "Measuring Marketing Productivity: Current Knowledge and Future Directions," *Journal of Marketing*, 68 (October), 76–89.

Santucci, Scott (2004), "Put Your Money Where the Value Is," *Productmarketing.com*, 2 (4), 22–26.

Schein, E.H. (1985), *Organizational Culture and Leadership*. San Francisco: Jossey-Bass.

Schmitt, P., K. Swope, and J. Walker (2000), "Collective Action with Incomplete Commitment: Experimental Evidence," *Southern Economic Journal*, 66 (4), 829–54.

Schotter, A. and B. Sopher (2003), "Social Learning and Convention Creation in Inter-Generational Games: An Experimental Study," *Journal of Political Economy*, 111 (3), 498–529.

——— and ——— (2007), "Advice and Behavior in Inter-Generational Ultimatum Games: An Experimental Approach," *Games and Economic Behavior*, 58 (2), 365–93.

Simon, Herbert A. and Craig A. Kaplan (1989), *Foundations of Cognitive Science*. Cambridge, MA: MIT Press.

Tannenbaum, S.I. and Gary Yukl (1992), "Training and Development in Work Organizations," *Annual Review of Psychology*, 43, 399–441.

VanBoskirk, Shar (2005), "Left Brain Marketing Planning," Forrester Research report, (May 16), 2–15.

The Wall Street Journal (2007), "Unraveling Rahodeb: A Grocer's Brash Style Takes Unhealthy Turn: Were Posts by Mackey, CEO of Whole Foods, a Case of Ethics, or Ego?" (July 20), A1.

Weick, K.E. (1987), "Organizational Culture as a Source of High Reliability" *California Management Review*, 29 (2), 112–27.

Wind, Yoram and Colin Crook (2005), *The Power of Impossible Thinking*. Upper Saddle River, NJ: Wharton School Publishing/Pearson Education.

Wright, W.F. and G.H. Bower (1992), "Mood Effects on Subjective Probability Assessment," *Organizational Behavior and Human Decision Processes*, 52 (2), 276–91.

Ravindra Chitturi, Rajagopal Raghunathan, & Vijay Mahajan

Delight by Design: The Role of Hedonic Versus Utilitarian Benefits

What is the relationship between product design benefits (hedonic versus utilitarian) and the postconsumption feelings of customer delight and satisfaction? The primary insights this research provides are as follows: (1) Products that meet or exceed customers' utilitarian needs and fulfill prevention goals enhance customer satisfaction (e.g., a car with antilock brakes and vehicle stability assist), and (2) products that meet or exceed customers' hedonic wants and fulfill promotion goals enhance customer delight (e.g., a car with panoramic sunroof and six-speaker audio system). Furthermore, the research finds that the primary antecedent feelings of satisfaction are the prevention emotions of confidence and security provided by utilitarian benefits, whereas the primary antecedent feelings of delight are the promotion emotions of cheerfulness and excitement provided by hedonic benefits. Finally, the results show that delighting customers improves customer loyalty, as measured by word of mouth and repurchase intentions, more than merely satisfying them. The authors discuss the theoretical contribution and strategic insights the research provides for product designers and marketers.

Keywords: design, delight, satisfaction, hedonic, utilitarian

For the past few decades, customer satisfaction has been the mantra by which firms have tried to manage the relationship between their products and customers. Much to their disappointment, managers are discovering that high levels of customer satisfaction do not necessarily translate into high levels of customer loyalty. For example, they might be surprised to find that more than 60% of customers who switch to another brand classify themselves as "satisfied" (Jones and Sasser 1995; Keiningham and Vavra 2001; Reichheld 1993). These findings point to a theoretically and substantively important research question: Why does customer satisfaction with products translate into such low levels of customer loyalty? Moreover, how can this be improved? A possibility is that customers want more than mere satisfaction; perhaps they want to be "delighted" in exchange for greater loyalty (Keiningham and Vavra 2001; Norman 2004; Rust and Oliver 2000). To address this possibility, we need to understand the nature of customers' consumption goals as they are fulfilled by the benefits offered by a product, their postconsumption emotional consequences, and their associated customer loyalty behavior as measured by word of mouth and repurchase intentions (Jacoby and Chestnut 1978).

This research builds on previous work on hedonic versus utilitarian benefits,[1] regulatory focus theory, customer satisfaction, and delight (e.g., Bagozzi, Gopinath, and Nyer 1999; Chitturi, Raghunathan, and Mahajan 2007; Oliver 1997; Rust and Oliver 2000) to address two questions: (1) Is there a difference in the nature of the emotional experience that results from the consumption of hedonic versus utilitarian benefits, and if so, do they differ in their ability to satisfy and/or delight the customer? and (2) Are customers who are delighted with product consumption more (or less) likely to indulge in word of mouth and report greater repurchase intentions than satisfied customers?

We address these questions by building on the following two well-established ideas: (1) Products evoke a positive emotional response when they exceed expectations on important benefit dimensions (Mittal, Ross, and Baldasare 1998; Oliver 1997), and (2) there is a fundamental difference in the type of goals hedonic and utilitarian benefits help attain (Chernev 2004; Chitturi, Raghunathan, and Mahajan 2007). Our contribution lies in qualifying these findings. We predict that the type of positive emotional response that product consumption evokes depends on whether the offering exceeds expectations on the utilitarian or the hedonic dimension. Specifically, we propose that whereas exceeding a utilitarian expectation merely evokes satisfaction, exceeding a hedonic expectation evokes delight. Furthermore, we propose that whereas failing to

Ravindra Chitturi is Assistant Professor of Marketing, Department of Marketing, College of Business and Economics, Lehigh University (e-mail: rac2@lehigh.edu). Rajagopal Raghunathan is Associate Professor of Marketing (e-mail: raj.raghunathan@mccombs.utexas.edu), and Vijay Mahajan is John P. Harbin Centennial Chair (e-mail: vijay.mahajan@mccombs.utexas.edu), Red McCombs School of Business, University of Texas at Austin. The first author acknowledges the support of the Laura Bonham Fund for Research at the University of Texas at Austin. The authors thank K. Sivakumar and James Maskulka for comments on a previous draft of this article. The authors also thank the three anonymous *JM* reviewers for their helpful suggestions.

---

[1] Voss, Spangenberg, and Grohmann (2003, p. 310, emphasis added) state, "Investigation of the *hedonic* and *utilitarian* components of attitude has been suggested in such diverse disciplines as sociology, psychology, and economics. This multidisciplinary recognition of the hedonic and utilitarian elements of *consumption* mirrors parallel theoretical development in marketing, mainly from a series of articles (e.g., Batra and Ahtola 1990; Chitturi, Raghunathan, and Mahajan 2007; Dhar and Wertenbroch 2000; Okada 2005)."

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (May 2008), 48–63

meet a utilitarian expectation evokes anger, failing to meet a hedonic expectation evokes dissatisfaction.

Findings across three studies with cell phones, laptop computers, and automobiles are consistent with our predictions. From a theoretical standpoint, our results indicate that the benefit dimension that produces delight may be different from that which produces satisfaction. From a substantive standpoint, our findings provide product designers and marketers with a better conceptual understanding of hedonic versus utilitarian product benefits and their relationship to delight, satisfaction, and loyalty.

We structure the rest of the article as follows: We begin with definitions of the terms "utilitarian," "hedonic," "promotion goals," and "prevention goals." We then discuss our theoretical model, in which we review relevant literature to derive our main predictions. We then provide empirical tests of these predictions and conclude with a discussion of our main theoretical and substantive contributions.

Theory and Hypotheses

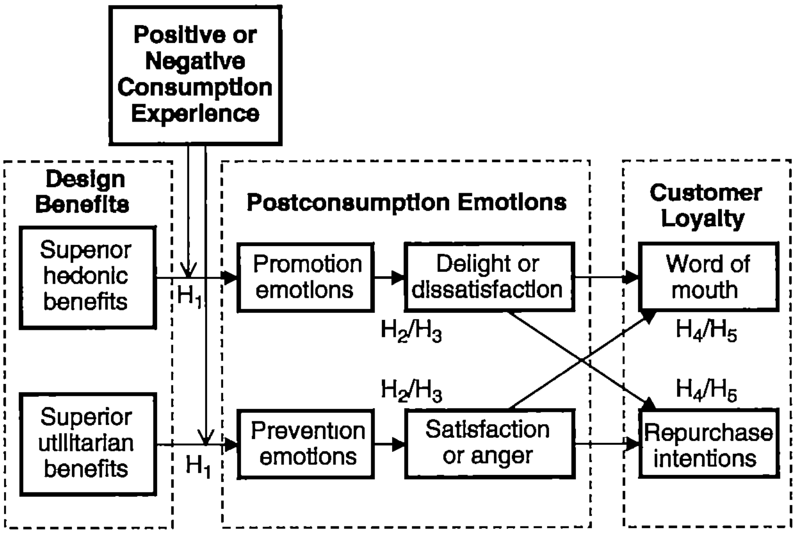
Definition of Utilitarian and Hedonic Benefits

Consistent with previous research in marketing, we use the term "utilitarian benefits" to refer to the functional, instrumental, and practical benefits of consumption offerings, and we use the term "hedonic benefits" to refer to their aesthetic, experiential, and enjoyment-related benefits (Batra and Ahtola 1990; Chitturi, Raghunathan, and Mahajan 2007; Dhar and Wertenbroch 2000; Strahilevitz and Myers 1998). Holbrook (1999, p. 8) notes that consumer value is an experience that results from the consumption of such benefits. In the context of cell phones, for example, the phone's battery life and sound volume are utilitarian benefits, whereas aesthetic appeal from its shape and color are hedonic benefits. Likewise, in the case of a car, fuel economy and safety are utilitarian benefits, whereas the sunroof and the luxurious interior are hedonic benefits.

Examining the relative consumer preference for these two benefit dimensions has been a fertile area of research in recent years (e.g., Chitturi, Raghunathan, and Mahajan 2007; Dhar and Wertenbroch 2000; Kivetz and Simonson 2002b; Okada 2005; Voss, Spangenberg, and Grohmann 2003). Prior research has focused primarily on assessing the relative weight that consumers attach to these two dimensions in preconsumption decisions. For example, Chitturi, Raghunathan, and Mahajan (2007) document that consumers attach greater importance to the hedonic (versus utilitarian) dimension, but only after a "necessary" level of functionality is met. Similarly, Kivetz and Simonson (2002a) document that consumers attach greater weight to the utilitarian (versus hedonic) dimension, unless they believe that they have "earned the right to indulge."

In contrast, we focus on the interplay between these two dimensions in influencing postconsumption emotional and behavioral consequences, an issue that needs more investigation. We build on the seminal work of Oliver and others (e.g., Mano and Oliver 1993; Oliver 1997; Rust and Oliver 2000; Westbrook 1987) to develop a conceptual framework of the role of hedonic and utilitarian benefits in evoking various types of positive and negative postconsumption emotions. Based on the recent work by Chernev (2004) and Chitturi, Raghunathan, and Mahajan (2007), the proposed framework (see Figure 1) posits that the consumption of a product's superior hedonic benefits evokes greater promotion emotions of cheerfulness and excitement, whereas the consumption of superior utilitarian benefits evokes greater prevention emotions of confidence and security. These promotion and prevention emotions are the primary antecedents of the postconsumption measures of delight and satisfaction, respectively, leading to greater word of mouth and repurchase intentions. This research empirically validates the proposed relationships in the model shown in Figure 1.

FIGURE 1
Proposed Conceptual Framework: Delight/Dissatisfaction Versus Satisfaction/Anger



Goals, Consumption Experience, and Emotional Consequences

Consumers usually purchase products on the basis of their expectations of the product's performance to meet their goals. Postconsumption emotions result from a perceived discrepancy between the expected performance and the actual performance (Oliver 1997). Specifically, depending on the level of expectancy confirmation or disconfirmation on important benefit dimensions, a product may evoke positive, negative, or a mixed set of emotions (Bagozzi, Gopinath, and Nyer 1999; Chaudhuri 2006; Westbrook and Oliver 1991). Extant research has implicitly assumed that expectancy confirmation or disconfirmation on one important benefit dimension is conceptually equivalent to expectancy confirmation or disconfirmation on any other equally important dimension; for example, prior research has implicitly assumed that a product that is twice as good as expected on the utilitarian dimension will produce the same type and intensity of positive emotional response as one that is twice as good as expected on the hedonic dimension. In contrast, as Figure 1 shows, we argue that expectancy confirmation (or disconfirmation) on the utilitarian dimension evokes a qualitatively different emotional response than does expectancy confirmation (or disconfirmation) on the hedonic dimension. This is primarily because the nature of the consumption experience and goals associated with utilitarian benefits are different from those associated with hedonic benefits (Chernev 2004; Chitturi, Raghunathan, and Mahajan 2007; Higgins 2001).

Prevention goals and utilitarian benefits, and promotion goals and hedonic benefits. There is converging evidence that the types of goals consumers expect to be fulfilled by the utilitarian dimension of a product are different from those they seek from the hedonic dimension (Chernev 2004). Specifically, whereas consumers expect the fulfillment of prevention goals on the utilitarian dimension, they expect the fulfillment of promotion goals on the hedonic dimension (e.g., Chernev 2004; Chitturi, Raghunathan, and Mahajan 2007; Higgins 1997, 2001). What are prevention and promotion goals?

According to the regulatory focus theory (Higgins 1997, 2001), prevention goals are those that ought to be met, such as "behaving in a safe and secure manner" and "being responsible." Fulfillment of prevention goals in the context of product consumption eliminates or significantly reduces the probability of a painful experience, thus making consumers experience emotions that result from fulfillment of prevention goals—for example, the positive prevention emotions of confidence and security (Chitturi, Raghunathan, and Mahajan 2007; Higgins 1997). For example, in the case of a car, prevention goals might be served by its safety features or the ten-year or 100,000-miles bumper-to-bumper warranty that the car manufacturer offers. Conversely, promotion goals are those that a person aspires to meet, such as "looking cool" or "being sophisticated." Fulfillment of promotion goals in the context of product consumption significantly increases the probability of a pleasurable experience, thus enabling consumers to experience emotions that result from the fulfillment of promotion goals—for example, the positive promotion emotions of cheerfulness and excitement (Chitturi, Raghunathan, and Mahajan 2007; Higgins 2001). The proposed conceptual framework in Figure 1 captures the relationships among hedonic versus utilitarian benefits, consumption experience, and promotion versus prevention emotions, respectively. However, do customers consider promotion and prevention goals equally, or do they consider one more important than the other? There is a large body of work in philosophy, sociology, and political science that assigns lower status to luxuries than to necessities in the hierarchy of needs (Berry 1994; Maslow 1970; Weber 1998). Moreover, Kivetz and Simonson (2002b) argue that luxuries are typically hedonic in nature and that necessities are typically utilitarian in terms of the benefits they offer to consumers. Holbrook (1999, p. 8) notes that consumer value is created by need- or want-satisfying experiences as a result of consuming such benefits. Furthermore, the principle of precedence positions the fulfillment of "needs" as a morally superior obligation than the fulfillment of "wants" (Berry 1994, p. 199). These findings collectively point to a conceptual parallel between luxuries–wants–hedonic benefits and necessities–needs–utilitarian benefits. Given that avoiding pain is a necessity (i.e., a need that a person must meet) and that seeking pleasure is a relative luxury (i.e., a want that a person aspires to meet), customers give higher priority to utilitarian benefits (i.e., fulfillment of prevention goals) than to hedonic benefits (i.e., fulfillment of promotion goals) (Chitturi, Raghunathan, and Mahajan 2007; Higgins 1997, 2001). However, when the prevention goals are met, the principle of hedonic dominance motivates customers to focus more on the continued fulfillment of promotion goals (Chitturi, Raghunathan, and Mahajan 2007). Therefore, the emotional consequences arising from meeting and failing to meet utilitarian and hedonic expectations differ significantly.

Utilitarian benefits and confidence/security, and hedonic benefits and excitement/cheerfulness. What are the positive emotions that result from the fulfillment of prevention and promotion goals? We propose that the type of positive emotion experienced depends on whether a utilitarian or a hedonic expectation is met.

To understand why this is so, consider the differences in the basic characteristics of the goals associated with the utilitarian and hedonic dimensions of product benefits we reviewed previously. The "must-meet" nature of prevention goals increases customers' focus on the utilitarian benefits of a product because it has been shown that utilitarian benefits are perceived as being closer to necessities or needs that help fulfill prevention goals (Chernev 2004; Higgins 1997; Kivetz and Simonson 2002b). However, when prevention goals are not fulfilled, customers experience increased pain in the form of negative feelings. For example, customer awareness that the absence of air bags and seat belts in a car could lead to severe injury in the event of a car accident is likely to evoke anxiety whenever the customer is in a fast-moving car. The presence of safety features, such as air bags, antilock brakes, and vehicle stability assist, reduces the pain of anxiety and increases feelings of security and confidence. Conversely, the "aspire-to-meet" nature of the promotion goals increases customers' focus on the hedonic benefits of a product (Chernev 2004). However, unlike the case of prevention goals, the nonfulfillment of promotion goals is perceived as a loss of pleasure rather than an increase in pain. It is because hedonic benefits are perceived as being closer to luxuries or wants that they fulfill promotion goals (Chernev 2004; Chitturi, Raghunathan, and Mahajan 2007; Higgins 1997; Kivetz and Simonson 2002b). A loss of pleasure is likely to evoke sadness and disappointment, but an increase in pain is likely to cause anger. For example, driving a convertible along the beautiful California coast on a sunny day is likely to be cheerful and exciting, leading to enhanced pleasure. However, driving a car without a convertible top is unlikely to be painful, though it could create a dissatisfactory experience. Therefore, consistent with Chernev (2004), Chitturi, Raghunathan, and Mahajan (2007), and Higgins (1997, 2001), the fulfillment of promotion goals by hedonic benefits is likely to evoke feelings of cheerfulness and excitement, and the fulfillment of prevention goals by utilitarian benefits is likely to evoke feelings of confidence and security.

> H_1: A positive consumption experience with utilitarian benefits evokes feelings of confidence and security, whereas a positive consumption experience with hedonic benefits evokes feelings of cheerfulness and excitement.

Utilitarian benefits and satisfaction, and hedonic benefits and delight. As we discussed previously, utilitarian benefits fulfill prevention goals and alleviate pain, and hedonic benefits fulfill promotion goals and enhance pleasure. However, to delight customers, it is not sufficient simply to avoid pain by meeting prevention goals. It is important that the fulfillment of prevention goals by the consumption of utilitarian benefits be complemented with the enhancement of pleasure by the consumption of hedonic benefits that fulfill promotion goals. For example, when the goals of Environmental Protection Agency mileage, passenger capacity, and reliability of an automobile are met, further improvements in mileage, passenger capacity, and/or reliability might lead to higher levels of satisfaction by increasing peace of mind, but not delight. The increase in peace of mind is due to a significant reduction in the probability of financial and physical pain from frequent refuelling, insufficient passenger capacity, and automobile breakdowns. However, for customers to be delighted, they need more than the low-arousal feelings associated with peace of mind from meeting or exceeding prevention goals. They need high-arousal pleasure from meeting or exceeding promotion goals. For example, automobile designers would need to offer hedonic benefits that help fulfill promotion goals, such as a convertible top, eight-speaker audio/video entertainment console, and beautiful interior. Therefore, consistent with the principle of hedonic dominance (Chitturi, Raghunathan, and Mahajan 2007), when prevention goals are fully met and consumers are satisfied, consumers give higher priority to hedonic benefits that help them meet promotion goals and enhance high-arousal feelings of excitement and cheerfulness, which in turn leads to delight. Chitturi, Raghunathan, and Mahajan (2007) show that at the time of purchase, customers feel more confident and secure with products that offer greater utilitarian benefits and feel more cheerful and excited with products that offer greater hedonic benefits. Therefore, customers associate the fulfillment of prevention goals with anticipated utilitarian benefits and the fulfillment of promotion goals with anticipated hedonic benefits of the products (Chernev 2004; Chitturi, Raghunathan, and Mahajan 2007). The degree to which these expectations of anticipated fulfillment of prevention and promotion goals materialize during consumption determines whether these preconsumption feelings of confidence, security, cheerfulness, and excitement will persist beyond consumption and contribute to customer satisfaction and/or delight. Therefore, we posit that products that meet or exceed utilitarian expectations will evoke low-arousal feelings of confidence and security, leading to satisfaction—an emotion that is also low in arousal (Hunt 1977; Oliver 1997, p. 300; Russell 1980). Conversely, products that meet or exceed hedonic expectations will evoke high-arousal feelings of excitement and cheerfulness, leading to delight—an emotion that is also high in arousal (for a review, see Ortony, Clore, and Collins 1988; see also Roseman 1991).

> H_2: A product that meets or exceeds utilitarian expectations evokes satisfaction, whereas one that meets or exceeds hedonic expectations evokes delight.

Prevention goals and anger, and promotion goals and dissatisfaction. Thus far, we have generated predictions about the types of positive emotions that are likely to be evoked by the fulfillment of prevention and promotion goals from utilitarian and hedonic benefits, respectively. A question that naturally follows is, What types of negative emotions are evoked in postconsumption contexts?

We know that consumers experience negative emotions when products fail to meet expectations (Mano and Oliver 1993; Westbrook 1987). We add the following qualification to this general rule: Whereas the failure to meet a utilitarian expectation leads to anger, the failure to meet a hedonic expectation merely leads to dissatisfaction. We evoke the must-meet and aspire-to-meet characteristics, respectively, of utilitarian and hedonic benefits to make these predictions (Higgins 1997).

That utilitarian goals are of a must-meet nature is simply another way of saying that consumers consider the fulfillment of utilitarian expectations a necessity (Chitturi, Raghunathan, and Mahajan 2007; see also Kivetz and Simonson 2002b). As such, when a product fails to meet a necessity (i.e., utilitarian expectation), consumers are likely to experience intense negative emotions that are high in arousal (Lazarus 1991; Roseman 1991). We expect this emotion to be anger related for two reasons. First, according to the precedence principle, consumers are likely to have paid particular attention to whether a product meets utilitarian requirements before selecting it. Any failure on the part of the product to meet utilitarian goals is more likely to be attributed to others (e.g., the manufacturer or retailer) than to themselves; appraisal theories of affect suggest that negative outcomes attributed to others are likely to lead to anger (Roseman 1991). Second, if consumers choose a more utilitarian product over a more attractive hedonic product to ensure fulfillment of prevention goals, there is already a sense of sadness and disappointment due to an anticipated loss of pleasure as a result of this trade-off (Chitturi, Raghunathan, and Mahajan 2007). Under such circumstances, if the utilitarian product does not fulfill minimum prevention goals, it is likely to add to the existing feeling of disappointment and evoke higher-arousal negative feelings. Given that the product did not perform as promised on the utilitarian dimension, there is a much stronger reaction and tendency to blame the manufacturer or retailer, leading to the high-arousal feeling of anger (Roseman 1991).

In contrast to the anger evoked by the failure to meet utilitarian expectations and, therefore, must-meet prevention goals, the failure to meet hedonic expectations and aspire-to-meet promotion goals is likely to evoke low-arousal feelings of sadness and disappointment, leading to dissatisfaction (Chitturi, Raghunathan, and Mahajan 2007; Inman, Dyer, and Jia 1997). This is because hedonic goals are perceived as a relative luxury compared with utilitarian goals, which are perceived as necessities (Chitturi, Raghunathan, and Mahajan 2007; see also Kivetz and Simonson 2002b). As such, products that do not meet hedonic expectations are likely to evoke negative emotions that are low in arousal, such as dissatisfaction.

> H_3: A product that fails to meet utilitarian expectations evokes anger, whereas one that fails to meet hedonic expectations evokes dissatisfaction.

Postconsumption Emotions, Word of Mouth, and Repurchase Intentions

An analysis of the emotional consequences of the fulfillment and frustration of utilitarian and hedonic goals is theoretically worthwhile. It is also substantively relevant because it helps predict postconsumption behavior. Because our primary goal is to understand how the promotion-focused and prevention-focused emotions influence delight, satisfaction, and customer loyalty, we examine how they influence two variables associated with loyalty: word of mouth and repurchase intentions (Jacoby and Chestnut 1978).

Satisfaction, delight, word of mouth, and repurchase intentions. According to psychoevolutionary theories of affect (e.g., Frijda 1987; Lazarus 1991; Plutchik 1980), different emotions are associated with different action tendencies (see Frijda 1987). For example, the action tendency associated with anger is one of "boiling inwardly," whereas that associated with sadness is one of feeling "helpless" (for a review, see Frijda 1986). Building on previous research (e.g., Mehrabian and Russell 1974), we argue that the likelihood that an action tendency will be translated into actual behavior depends on the level of arousal associated with the emotion in question. The feelings of cheerfulness and excitement arising from the fulfillment of promotion goals by hedonic benefits are high-arousal feelings (Lazarus 1991; Roseman 1991). They are also the antecedent feelings of delight, an emotion accompanied by high arousal. Conversely, the fulfillment of prevention goals by utilitarian benefits leads to the low-arousal feelings of confidence and security (Higgins 1997; Lazarus 1991). Moreover, the feelings of confidence and security are antecedent feelings of satisfaction, an emotion accompanied by relatively lower arousal compared with delight (Oliver 1997). Therefore, we hypothesize the following:

> H_4: Consumers are more likely to indulge in (a) positive word of mouth and (b) repeat purchase behavior when their hedonic expectations are met or exceeded than when their utilitarian expectations are met or exceeded.

Anger, dissatisfaction, repurchase intentions, and word of mouth. Thus far, we have argued that delighted (versus satisfied) customers are more likely to (1) express positive word of mouth and (2) indicate intentions of repeat purchase. What happens when the consumption experience is negative?

As we discussed previously, there is a conceptual parallel between necessities–needs–utilitarian benefits and luxuries–wants–hedonic benefits (Chernev 2004; Chitturi, Raghunathan, and Mahajan 2007; Higgins 1997; Kivetz and Simonson 2002b). In addition, Maslow's hierarchy of needs and the precedence principle both position necessities–needs–utilitarian benefits as of greater importance than luxuries–wants–hedonic benefits. Chitturi, Raghunathan, and Mahajan (2007) show that consumers are focused more on the utilitarian benefits than on the hedonic benefits of a product until their minimum expectations of fulfilling prevention goals are met. Furthermore, a utilitarian benefit is a promise of a certain level of functionality by the manufacturer or the retailer. When this promise is not fulfilled, customers blame the retailer and/or the manufacturer. When negative feelings are attributable to an entity, customers feel angry (Lazarus 1991; Roseman 1991). However, in the case of hedonic benefits, such as style and visual appeal, "what you see is what you get." The customer, not the manufacturer, determines at the time of purchase whether the product is stylish and attractive. Under such circumstances, customers are more likely to blame themselves than the manufacturer if their friends do not find the product stylish and attractive. Therefore, not meeting minimum utilitarian expectations of functionality generates a much more intense negative feeling, such as anger, than a less intense feeling, such as disappointment or dissatisfaction that results from not meeting minimum hedonic expectations. Because anger is accompanied by higher levels of arousal than dissatisfaction, we expect that the failure of products to meet utilitarian (versus hedonic) expectations leads to greater negative word of mouth and lower repurchase intentions.

> H_5: Consumers are (a) more likely to indulge in negative word of mouth and (b) less likely to engage in repeat purchase when their utilitarian expectations are not met than when their hedonic expectations are not met.

Overview of Studies and Results

In three studies with cell phones, laptop computers, and cars, we show that the hedonic and utilitarian benefits of a product differ in their ability to delight and satisfy customers. We chose the products to ensure that participants would be familiar with their attributes and benefits and could imagine the products in various usage scenarios. At the same time, we wanted the chosen products to be both visible and useful to enhance the relevance of both the hedonic and the utilitarian benefits. The first two studies with cell phones and laptop computers involved student participants at a North American university for extra credit. We conducted the third study with automobile owners to tap into real feelings based on real consumption experience.

Study 1

Design and Task

We used a 2 (product design benefits[2]: hedonic versus utilitarian) × 4 (consumption experience: confirm/confirm, confirm/disconfirm, disconfirm/confirm, disconfirm/disconfirm) between-subjects design. A total of 240 undergraduate students participated for extra credit. The participants were randomly assigned to each of the eight cells. Participants were given a booklet titled "Consumer Decision Making Questionnaire," with the following starting instructions on the first page:

> In this questionnaire, we are interested in finding how you feel about a product after its purchase and consumption. In the following pages, you will read about two scenarios that describe your needs, attributes of a chosen product, and subsequent consumption experience. Please read the two scenarios carefully and respond to the questions that follow.

On the following page, participants read information about the two cell phone choices. Each cell phone was described as a combination of the three hedonic and three utilitarian attributes. Three of the attributes were primarily hedonic and added to the hedonic consumption benefits, and the other three attributes were primarily utilitarian and added to the utilitarian consumption benefits.[3] We combined the three hedonic and three utilitarian attributes with a picture of the cell phone. Each of the six attributes and the cell phone picture had two levels and associated consumption benefits—high and medium.[4] The group of attributes offering benefits—hedonic or utilitarian—labeled as "high" was given a *Consumer Reports* rating of 4.5 out of 5, and the group labeled as "medium" was given a *Consumer Reports* rating of 3.0 out of 5. After reviewing the information about each cell phone including the picture, participants read two product choice and consumption scenarios. The first scenario described the purchase and subsequent consumption experience with one of the cell phones, and the second scenario described the purchase and subsequent consumption experience with the other cell phone. The consumption experience within a group condition was the same across the two scenarios. Only the cell phone choice (hedonic versus utilitarian) changed between the two scenarios. The participants were then told to imagine themselves in the first scenario, and they answered questions about how they would feel in that situation. They reported their level of feelings on a total of 14 anchor measures of discrete positive and negative emotions.[5] They rated 14 emotions (i.e., guilt, anxiety, sadness, dissatisfaction, regret, anger, disappointment, surprise, security, confidence, excitement, satisfaction, cheerfulness, and delight) with the following instructions: "Based on the overall experience of using my current product, I feel ..." (seven-point scale anchored by 1 = "not at all" and 7 = "extremely"; for sample measures, see Appendix A). This was followed by a four-item scale that measured arousal (Mehrabian and Russell 1974). The four-item scale included measures comprising stimulated/relaxed, excited/calm, aroused/unaroused, and jittery/dull. Participants then answered questions about positive word of mouth, negative word of mouth, and repurchase intentions (for sample measures, see Appendix A). These measures of emotions, word of mouth, and repurchase intentions served as dependent variables across the treatment conditions. Finally, participants completed the section containing manipulation checks and demographic variables.

Stimuli Construction

The objective was to construct more realistic stimuli while retaining the level of control that is needed to test the hypotheses. To accomplish this objective, we conducted two pretests. For Pretest 1, we made a comprehensive list of attributes based on real product manuals of some of the most popular cell phones in the market (e.g., Nokia, Samsung, Motorola). This led to a list of more than 50 attributes. For practical reasons, the goal was to identify the top 3 most influential attributes that offer hedonic benefits and the top 3 most influential attributes that offer utilitarian benefits to construct the stimuli for the experiment. To accomplish this, we asked more than 100 participants to rate all the attributes in terms of importance and impact on purchase decision. On the basis of the Pretest 1 responses, we selected the top 15 most influential attributes on purchase decision. To test our theory, we needed to construct product stimuli that had a group of attributes that was primarily hedonic and another group that was primarily utilitarian in terms of the benefits offered. On the basis of the ratings, we selected 3 attributes to construct the hedonic dimension and another 3 to construct the utilitarian dimension of the product benefits. We manipulated each attribute to create two levels. We then combined these attributes with the pretested pictures of cell phones to create the stimuli. One of the cell phone choices was highly stylish and attractive (hedonic) with a medium *Consumer Reports* rating of 3.0 out of 5 on functionality (utilitarian). The other was a highly functional (utilitarian) cell phone with a medium *Consumer Reports* rating of 3.0 out of 5 on style and attractiveness (hedonics).

---

[2] The benefits offered by the attributes of a product can be broadly categorized along two dimensions: hedonic and utilitarian. These benefits can be high or low, depending on the product (Crowley, Spangenberg, and Hughes 1992). Consistent with the work of Dhar and Wertenbroch (2000) and Okada (2005), this research conceptualizes hedonic (or utilitarian) product alternatives as ones that offer relatively superior hedonic (or utilitarian) benefits.

[3] Using the HED/UTI scale that Voss, Spangenberg, and Grohmann (2003) developed, we selected the attributes for the primary nature of their benefits—hedonic or utilitarian. We selected three attributes to construct the hedonic dimension and another three to construct the utilitarian dimension of the product benefits.

[4] We pretested cell phone and laptop computer pictures for their style and attractiveness. A collection of nine cell phone and laptop computer pictures were rated by 47 students on a scale from 1 ("low") to 10 ("high") on the overall measure of "style and attractiveness." We picked the picture that was rated the highest (cell phone mean = 8.7; laptop computer mean = 7.5) and the one that was rated in the middle (cell phone mean = 4.6; laptop computer mean = 4.1) to construct the stimuli. We used the most stylish and attractive picture with superior hedonic benefits combined with medium utilitarian benefits and the other picture with medium hedonic benefits combined with superior utilitarian benefits to construct the stimuli for the two products used in Study 1 and Study 2.

[5] Single-item discrete emotion measures are prevalent in emotion literature. For a more detailed discussion, see Larsen and Fredrickson (1999, pp. 40–60). For a more detailed discussion on the effectiveness of single-item versus multi-item measures, see Bergkvist and Rossiter (2007).

Stimuli Description

The utilitarian benefits dimension of each cell phone comprised the level of network coverage (98% versus 95%), battery capacity (three days versus two days), and sound clarity (high versus medium). The hedonic benefits dimension comprised the oyster flip phone feature (yes/no), the option to change phone colors (yes/no), and the ability to program new ring tones (yes/no). These attribute descriptions and levels were combined with pictures of the two cell phones. One picture showed a highly stylish and attractive oyster flip phone, and the other showed a medium-rated nonflip feature cell phone.

Manipulation Checks

Expectancy disconfirmation. On a seven-point scale, participants indicated the extent to which the consumption experience met their expectations on the hedonic dimension (1 = "better than expected," and 7 = "worse than expected"). We repeated the same question for the utilitarian dimension. We reverse-coded the measures for data analysis. The t-tests showed that the manipulation was successful. The means for the four experience conditions for the product with superior hedonic benefits were (1) confirm/confirm = 5.84/4.96, (2) disconfirm/disconfirm = 2.8/2.32, (3) confirm/disconfirm = 5.18/2.43, and (4) disconfirm/confirm = 3.1/4.61. Similarly, the means for the four experience conditions for the product with superior utilitarian benefits were (1) confirm/confirm = 5.27/5.56, (2) disconfirm/disconfirm = 2.73/2.51, (3) confirm/disconfirm = 4.98/2.77, and (4) disconfirm/confirm = 2.4/5.11.

Results and Discussion

To test H_1, we ran a multivariate analysis of variance (MANOVA) across four experience conditions and two product types (superior hedonic benefits versus superior utilitarian benefits) for prevention emotions of confidence and security and promotion emotions of excitement and cheerfulness. The results from the MANOVA were significant for product type (Wilks' $\lambda = .605$, $F = 37.389$, $p < .01$) and experience (Wilks' $\lambda = .510$, $F = 14.69$, $p < .01$), and the univariate tests for cheerfulness, excitement, confidence, and security were significant as well ($p < .05$). The cell means and standard deviations for all the positive emotions as a result of consumption experience with a cell phone with superior hedonic benefits versus superior utilitarian benefits appear in Table 1. A two-sample t-test found that in the case of a positive consumption experience, consumers feel greater intensity of excitement and cheerfulness with a more hedonic cell phone than with a more utilitarian cell phone ($t_{excitement} = 6.84$, $p < .01$; $t_{cheerfulness} = 3.38$, $p < .01$; see Appendix B, Panel 1). Conversely, positive experience with a utilitarian cell phone leads to greater security and confidence ($t_{security} = 5.32$, $p < .01$; $t_{confidence} = 5.22$, $p < .01$, see Appendix B, Panel 1). The result with cell phones is consistent with our predictions in H_1. To test H_2 and H_3, we ran a MANOVA across two product types for the emotions of delight, satisfaction, anger, and dissatisfaction. The results for the two product-type groups were significant (Wilks' $\lambda = .903$, $F = 12.518$, $p < .01$), and the univariate tests for delight, satisfaction, anger, and dissatisfaction were significant. To show that delight and satisfaction are affected differently by hedonic and utilitarian benefits, we

TABLE 1
Postconsumption Emotional Responses for Positive Experience with Hedonic Versus Utilitarian Product Benefits

Postconsumption Customer Emotions	Product with Superior Hedonic Benefits: M (SD)	Product with Superior Utilitarian Benefits: M (SD)	Estimate of Difference: t-Value
	Cell Phone (Study 1)		
Excitement	6.07 (1.31)	3.70 (1.37)	6.84*
Cheerfulness	5.47 (1.36)	4.20 (1.54)	3.38*
Delight	5.87 (1.31)	4.47 (1.53)	3.82*
Security	3.80 (1.54)	5.70 (1.21)	5.32*
Confidence	3.47 (1.50)	5.47 (1.41)	5.22*
	Laptop Computer (Study 2)		
Excitement	5.83 (1.36)	3.50 (1.39)	6.57*
Cheerfulness	5.03 (1.40)	3.80 (1.27)	3.57*
Delight	5.57 (1.52)	4.17 (1.32)	3.81*
Security	3.97 (1.61)	5.13 (1.31)	3.08*
Confidence	3.77 (1.30)	5.03 (1.81)	3.11*

*Significant at $p < .01$

also tested the difference between delight and satisfaction across hedonic and utilitarian cell phones and found this difference to be significant. As we discussed previously, the changes in benefits offered by the cell phone (hedonic versus utilitarian) have a greater influence on the emotion of delight than on satisfaction. The results show that changes in customers' postconsumption feeling of delight are much greater between the hedonic and the utilitarian cell phone than the changes in satisfaction for the same two cell phones ($t_{\text{Diff-of-Diff}} = 2.48$, $p < .01$). Furthermore, the tests show that changes in customers' postconsumption feeling of anger are much greater between the hedonic and the utilitarian cell phone than the changes in the feeling of dissatisfaction for the same two cell phones ($t_{\text{Diff-of-Diff}} = 2.02$, $p < .05$). The findings about customer emotions show that the results are significant and in the right direction; hedonic cell phones are better at fulfilling promotion-focused emotional goals, and utilitarian cell phones are better at fulfilling prevention-focused emotional goals (Chernev 2004; Higgins 2001).

Next, in terms of H_4 and H_5, we want to find out whether differences in hedonic versus utilitarian product benefits and the emotions elicited in the consumption experience are also accompanied by significantly different levels of arousal, positive word of mouth, negative word of mouth, and repurchase intentions in the case of positive and negative consumption experiences. The two-sample t-test shows that customers experience greater arousal with the hedonic cell phone than with the utilitarian cell phone ($t_{\text{arousal}} = 4.10$, $p < .01$). The two-sample t-test shows that customers are more likely to indulge in positive word-of-mouth behavior and have greater repurchase intentions in the case of positive consumption experience with a hedonic cell phone than with a utilitarian cell phone ($t_{\text{WOM}} = 5.19$, $p < .01$; $t_{\text{RPI}} = 4.73$, $p < .01$). Conversely, consumers experience greater arousal with a utilitarian cell phone than with a hedonic one in the case of negative consumption experience ($t_{\text{arousal}} = 3.54$, $p < .01$). Therefore, as we expected, consumers are more likely to indulge in negative word-of-mouth behavior and have less repurchase intentions with a utilitarian cell phone than with a hedonic cell phone ($t_{\text{NWOM}} = 4.90$, $p < .01$; $t_{\text{LRPI}} = 3.70$, $p < .01$). This demonstrates that the consumption of hedonic versus utilitarian benefits offered by a product indeed generates significantly different intensities of postconsumption emotions and is accompanied by higher levels of positive (or negative) word of mouth and repurchase intentions when the experience is positive (or negative).

Study 2

Design and Task

In Study 1, we used cell phones to collect our data. In Study 2, we used laptop computers to test our hypotheses. On the basis of a pretest for perceived importance of hedonic and utilitarian benefits with the cell phones and laptop computers, we found that participants consider utilitarian benefits relatively more important in the case of laptop computers and give relatively more weight to hedonic benefits in the case of cell phones. Therefore, this study tests the same hypotheses with a relatively more utilitarian product category (i.e., laptop computers) than Study 1. The design was identical to Study 1. Another group of 240 undergraduate business students participated for extra credit. We assigned them randomly to each of the eight cells, and the instructions were the same as in Study 1. We used the dependent measures of emotion, word of mouth, and repurchase intentions, followed by manipulation checks for expectancy and needs disconfirmation, as well as demographics.

Stimuli Description

We followed all the same steps and processes to construct the laptop computer stimuli for Study 2 that we followed to construct the cell phone stimuli for Study 1. After two pretests, we used two laptop computers that were a combination of three attributes for the hedonic dimension and another three attributes for the utilitarian dimension of the consumption benefits. The utilitarian dimension comprised the level of processing speed (3 GHz versus 1.5 GHz), memory size (20 GB versus 10 GB), and audio clarity (high versus medium). The hedonic dimension comprised screen size (wide screen versus regular screen), change laptop color (yes/no), and weight (five pounds versus ten pounds). We combined these attribute descriptions and levels with pictures of the two laptop computers. One picture showed a highly stylish and attractive laptop computer, and the other showed a medium-rated laptop computer.

Results and Discussion

To test H_1, we ran a MANOVA across four experience conditions and two product types (superior hedonic benefits versus superior utilitarian benefits) for the prevention emotions of confidence and security and for the promotion emotions of excitement and cheerfulness. The results from the MANOVA were significant for product type (Wilks' $\lambda = .721$, $F = 33.261$, $p < .01$) and experience (Wilks' $\lambda = .498$, $F = 16.39$, $p < .01$), and the univariate tests for all measures were significant as well ($p < .05$). A two-sample t-test found that in the case of a positive consumption experience, consumers feel greater intensity of excitement and cheerfulness with a hedonic laptop computer than with a utilitarian one ($t_{\text{excitement}} = 6.57$, $p < .01$; $t_{\text{cheerfulness}} = 3.57$, $p < .01$; for cell means and standard deviation statistics, see Table 1). Conversely, positive experience with a utilitarian laptop computer leads to greater security and confidence ($t_{\text{security}} = 3.08$, $p < .01$; $t_{\text{confidence}} = 3.11$, $p < .01$). The result with laptop computers is consistent with the predictions in H_1. To test H_2 and H_3, we ran a MANOVA across two product types for the emotions of delight, satisfaction, anger, and dissatisfaction. The results for the two product-type groups were significant (Wilks' $\lambda = .843$, $F = 17.321$, $p < .01$), and the univariate tests for delight, satisfaction, anger, and dissatisfaction were significant. Consistent with Study 1, the results show that changes in customers' postconsumption feeling of delight are much greater between hedonic and utilitarian laptop computers than the changes in feeling of satisfaction for the same two laptop computers ($t_{\text{Diff-of-Diff}} = 2.80$, $p < .01$). Furthermore, the tests show that changes in customers' postconsumption feeling of anger are much greater between the hedonic and the utilitarian laptop computer than the changes in feeling of dissatisfaction for the same two laptop computers ($t_{\text{Diff-of-Diff}} = 2.12$, $p < .05$; for means and standard deviations, see Table 2).

Next, we tested H_4 and H_5. The results are consistent with the results from Study 1. The two-sample t-test shows that customers experience greater arousal with a hedonic laptop computer than with a utilitarian laptop computer ($t_{\text{arousal}} = 3.84$, $p < .01$). The two-sample t-test shows that customers are more likely to indulge in positive word-of-mouth behavior and have greater repurchase intentions in the case of positive consumption experience with a hedonic laptop computer than with a utilitarian laptop computer ($t_{\text{WOM}} = 3.90$, $p < .01$; $t_{\text{RPI}} = 4.00$, $p < .01$). Conversely, consumers experience greater arousal with a utilitarian laptop computer than with a hedonic one in the case of negative consumption experience ($t_{\text{arousal}} = 3.54$, $p < .01$). Therefore, as we expected, consumers are more likely to indulge in negative word-of-mouth behavior and have lower repurchase intentions with a utilitarian laptop computer than with a hedonic laptop computer ($t_{\text{NWOM}} = 3.77$, $p < .01$; $t_{\text{LRPI}} = 2.89$, $p < .01$). The results are consistent across the two studies with cell phones and laptop computers.[6]

Study 3

We conducted Studies 1 and 2 with undergraduate students, and the experimental design measured consumption emotions and loyalty measures on the basis of "imagined experience scenarios" rather than "real experiences." To address this issue and to improve the external validity, we conducted Study 3 with car owners who visited national automobile service franchises. We measured the same set of 14 dependent measures of postconsumption emotions and customer loyalty on the basis of owners' real experiences with their car. The results are consistent with the findings in Studies 1 and 2 and provide additional insights into our research question.

Design and Task

We designed the questionnaire to collect information about car owners' feelings about their car based on their consumption experience. A total of 142 car owners participated in the study. They rated 14 emotions (i.e., guilt, anxiety, sadness, dissatisfaction, regret, anger, disappointment, surprise, security, confidence, excitement, satisfaction, cheerfulness, and delight) in response to the statement, "Based on the overall experience of using my current car, I feel …" on a seven-point scale ranging from 1 = "not at all" to 7 = "extremely" (for sample measures, see Appendix A). After the emotion measures, the participants rated how likely they would be to recommend their car to others as a measure of the positive word of mouth and how likely they would be to repurchase this car if they were to do it all over again. We measured word of mouth and repurchase intentions on a seven-point scale ranging from 1 = "not at all likely" to 7 = "extremely likely." This was followed by participants' overall impression of the product on the hedonic dimension of "style and attractiveness" and the utilitarian dimension of "functionality" on a seven-point scale ranging from 1 = "not at all impressive" to 7 = "extremely impressive," as well as by a four-item measure of arousal (Mehrabian and Russell 1974). Participants then answered questions about whether the product met their expectations on the "style and attractiveness" and "functionality" dimensions. Finally, participants provided demographic information.

Results and Discussion

On the basis of the median split of participants' responses on the overall rating of "style and attractiveness" and "functionality" of their cars, we created four groups (i.e., 1 = high/high, 2 = low/low, 3 = high/low, and 4 = low/high). The results for emotions from a MANOVA were significant for the four group conditions (Wilks' $\lambda = .444$, $F = 1.89$, $p < .01$), and the univariate tests for all emotions except guilt, anxiety, and surprise were significant as well ($p < .05$). Studies in the literature have linked pleasant surprise (i.e., positive disconfirmation) as a necessary condition for delight, but we did not find that to be the case in this study (for more elaboration on the relationship between surprise and delight, see the "General Discussion" section). Our pri-

TABLE 2
Postconsumption Emotional Responses for Negative Experience with Hedonic Versus Utilitarian Product Benefits

Postconsumption Customer Emotions	Product with Superior Hedonic Benefits: M (SD)	Product with Superior Utilitarian Benefits: M (SD)	Estimate of Difference: t-Value
	Cell Phone (Study 1)		
Dissatisfaction	6.20 (1.10)	5.27 (1.62)	2.62*
Anger	3.70 (1.78)	5.73 (1.01)	5.07**
	Laptop Computer (Study 2)		
Dissatisfaction	5.97 (1.13)	4.80 (1.40)	3.55**
Anger	3.80 (1.65)	5.80 (1.06)	5.58**

*Significant at $p < .05$.
**Significant at $p <$ 01

---

[6] We replicated a study with the same design with nonstudents using cell phones. The results are consistent with the first two studies with students. Therefore, in the interest of space, we do not report the results of that study.

mary focus in this article is to identify the differences in the antecedents of delight and satisfaction resulting from the respective consumption of products with superior hedonic benefits and superior utilitarian benefits. In line with the work of Higgins (1997, 2001) and Chitturi, Raghunathan, and Mahajan (2007), we group the emotions of confidence and security by averaging them to form a measure of prevention emotions and by averaging the emotions of excitement and cheerfulness to form a measure of promotion emotions. As Study 1 shows, a positive consumption experience with superior hedonic benefits results in promotion emotions, and a positive consumption experience with superior utilitarian benefits results in prevention emotions.

We propose and test two models to study the relationship between hedonic and utilitarian benefits with loyalty, as measured by word of mouth and repurchase intentions (see Figure 2). Finally, we conclude with the test of a combined model to assess the relative influence of hedonic and utilitarian benefits on delight, satisfaction, word of mouth, and repurchase intention. The means, standard deviations, and correlation matrix for the relevant measures appear in Table 3. Although delight and satisfaction are positively correlated ($\alpha = .52$), we also find evidence that there are significant differences between delight and satisfaction in terms of the antecedent emotions. The discussion of the results of our analysis follows.

Model 1 (see Figure 2) tests the proposed relationships among promotion emotions, delight, and loyalty, as measured by word of mouth and repurchase intentions, using mediation analysis (Baron and Kenny 1986). Model 2 (see Figure 2) tests the proposed relationships among prevention emotions, satisfaction, word of mouth, and repurchase intentions. The results from testing Model 1 reveal that delight partially mediates the relationship between promotion emotions and word of mouth (i.e., both promotion emotions and delight significantly influence word-of-mouth behavior). Furthermore, consistent with Oliver, Rust, and Varki's (1997) findings, delight does not mediate (fully or partially) the relationship between promotion emotions and repurchase intentions; the relationship between delight and repurchase intentions becomes nonsignificant when regressed with promotion emotions ($\beta_{\text{delight}} = .22$, $t = 1.12$, $p < .227$) (for path coefficient values and significance level, see Figure 2, Model 1). The results from testing Model 2 reveal that satisfaction partially mediates the relationship between prevention emotions and word of mouth, as well as repurchase intentions (for path coefficient values and significance level, see Figure 2, Model 2). To supplement the tests for H_{4a} and H_{4b} from Studies 1 and 2, which show that hedonic benefits influence delight more than satisfaction and that utilitarian benefits influence satisfaction more than delight, we combine the two models shown in Figure 2. A summary of the relevant results of regression with a combined model appears in Table 4. The results show that both delight and satisfaction directly influence word of mouth ($\beta_{\text{delight}} = .163$, $p < .05$; $\beta_{\text{satisfaction}} = .546$, $p < .000$; adjusted $R^2 = .374$), and delight significantly improves word of mouth in addition to satisfaction. Similarly, both delight and satisfaction directly influence repurchase intentions ($\beta_{\text{delight}} = .243$, $p < .034$; $\beta_{\text{satisfaction}} = .714$, $p < .000$; adjusted $R^2 = .371$), and delight significantly improves repurchase intentions in addition to satisfaction. The results demonstrate that customer delight improves customer loyalty by improving word of mouth and repurchase intentions. However, what about the role of hedonic versus utilitarian design benefits in delighting versus satisfying a customer? The results in Table 4 show that only hedonic benefits contribute to delight. The effect of utilitarian benefits on customer delight is not significant ($\beta_{\text{utilitarian}} = .124$, n.s.; $\beta_{\text{hedonic}} = .405$, $p < .000$; adjusted $R^2 = .170$). However, in the case of satisfaction, both hedonic and utilitarian benefits significantly improve satisfaction, and consistent with the results from Studies 1 and 2, utilitarian benefits have greater influence on satisfaction than hedonic benefits ($\beta_{\text{utilitarian}} = .298$, $p < .009$; $\beta_{\text{hedonic}} = .197$, $p < .019$; adjusted $R^2 = .180$). Further analysis of the relationship among design benefits, promotion emotions, and prevention emotions reveals that only hedonic benefits influence promotion emotions ($\beta_{\text{utilitarian}} = .070$, n.s.; $\beta_{\text{hedonic}} = .478$, $p < .000$; adjusted $R^2 = .211$), whereas both hedonic and utilitarian benefits influence prevention emotions. As we predicted, utilitarian benefits have a greater influence on prevention emotions than hedonic benefits ($\beta_{\text{utilitarian}} = .400$, $p < .000$; $\beta_{\text{hedonic}} = .227$, $p < .019$; adjusted $R^2 = .271$). Consistent with our proposed model, we find that delight is influenced only by promotion emotions ($\beta_{\text{prevention emotions}} = .126$, n.s.; $\beta_{\text{promotion emotions}} = .797$, $p < .000$; adjusted $R^2 = .741$), whereas satisfaction is influenced by both promotion and prevention emotions ($\beta_{\text{prevention emotions}} = .373$, $p < .000$; $\beta_{\text{promotion emotions}} = .325$, $p < .000$; adjusted $R^2 = .399$). As we discussed previously and summarize in Table 4, the results show a stronger association of hedonic benefits with promotion emotions and delight and a stronger association of utilitarian benefits with prevention emotions and satisfaction.

General Discussion

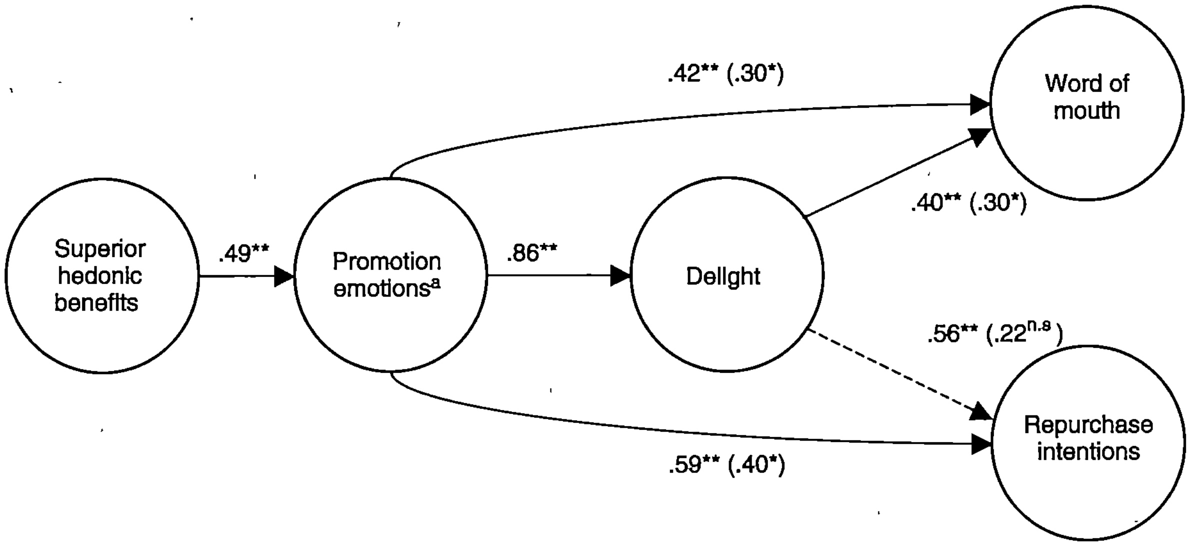
Should firms try to delight the customer? What is the difference between delight and satisfaction? These are some of the questions that have been explored in prior research (Keiningham and Vavra 2001; Oliver, Rust, and Varki 1997; Rust and Oliver 2000). In this article, we studied and found evidence for the research proposition that the type and the intensity of the emotional experience arising from the consumption of hedonic benefits are qualitatively different from those of utilitarian benefits. This difference in emotional experience results in significantly different levels of customer delight, satisfaction, word of mouth, and repurchase intentions. We explored this question with studies involving cell phones, laptop computers, and automobiles. Studies 1 and 2 involved students from a North American university, and Study 3 involved automobile owners and tapped into their consumption experiences with cars. The results across the three studies were consistent and support the fundamental research proposition and the proposed conceptual framework (see Figure 1).

The primary insights this research provides are as follows: (1) Products that meet or exceed customers' utilitarian needs and fulfill prevention goals enhance customer satis-

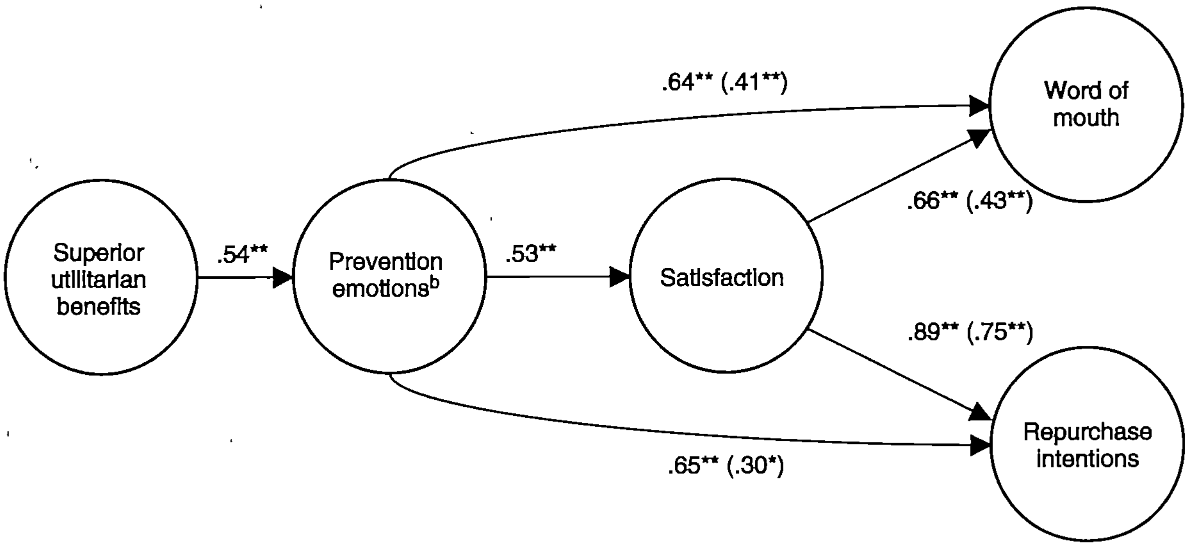
FIGURE 2

Models of Postconsumption Positive Emotions and Loyalty of Car Owners Based on the Consumption of Superior Hedonic Benefits (Model 1) and Superior Utilitarian Benefits (Model 2)

Model 1: Promotion Emotions, Delight, and Loyalty (Study 3)



Model 2: Prevention Emotions, Satisfaction, and Loyalty (Study 3)



*Relationship is significant at $p < .05$.

**Relationship is significant at $p < .01$.

[a] Promotion emotions of excitement and cheerfulness (Higgins 1997, 2001).

[b] Prevention emotions of confidence and security (Higgins 1997, 2001).

faction, and (2) products that meet or exceed customers' hedonic wants and fulfill promotion goals enhance customer delight. Furthermore, the research finds that the primary antecedent feelings of satisfaction are the prevention emotions of confidence and security provided by utilitarian benefits (Chitturi, Raghunathan, and Mahajan 2007; Higgins 1997, 2001), whereas the primary antecedent feelings of delight are the promotion emotions of cheerfulness and excitement provided by hedonic benefits (Chitturi, Raghunathan, and Mahajan 2007; Higgins 1997, 2001). In addition, the research shows that customer delight improves customer loyalty by improving both word of mouth and repurchase intentions.

Theoretical Contribution

Building on the work of Higgins (1997, 2001), Chernev (2004), and Chitturi, Raghunathan, and Mahajan (2007), we proposed that the goals of a prevention focus served by utili-

TABLE 3
Postconsumption Positive Emotions of Car Owners (Study 3): Means, Standard Deviations, and Correlations

	M (SD)	1	2	3	4	5	6	7	8
1. Delighted	4.53 (1.70)	1							
2. Satisfied	5.35 (1.31)	.52	1						
3. Excited	4.58 (1.85)	.78	.58	1					
4. Confident	5.09 (1.38)	.63	.61	.64	1				
5. Cheerful	4.62 (1.67)	.86	.57	.84	.65	1			
6. Secure	5.11 (1.48)	.52	.44	.48	.81	.56	1		
7. Word of mouth	5.28 (1.47)	.57	.59	.48	.62	.44	.51	1	
8. Repurchase intentions	4.67 (1.99)	.48	.59	.49	.47	.47	.38	.78	1

Notes: All correlations are statistically significant ($p < .01$).

TABLE 4
Full Model Statistics of Relationships Involving Hedonic and Utilitarian Design Benefits, Delight, Satisfaction, Word of Mouth, and Repurchase Intentions (Study 3)

Dependent Variable	Independent Variable	Standardized Coefficient β	p-Value	Adjusted R^2
Word of mouth	Delight	.163	.050	.374
	Satisfaction	.546	.000	
Repurchase intentions	Delight	.243	.034	.371
	Satisfaction	.714	.000	
Delight	Hedonic benefits	.405	.000	.170
	Utilitarian benefits	.124	n.s.	
Satisfaction	Hedonic benefits	.197	.019	.180
	Utilitarian benefits	.298	.009	
Promotion emotions	Hedonic benefits	.478	.000	.211
	Utilitarian benefits	.070	n.s.	
Prevention emotions	Hedonic benefits	.227	.006	.271
	Utilitarian benefits	.400	.000	
Delight	Promotion emotions^a	.797	.000	.741
	Prevention emotions^b	.126	n.s.	
Satisfaction	Promotion emotions	.325	.000	.399
	Prevention emotions	.373	.000	
Delight	Promotion emotions	.765	.000	.742
	Prevention emotions	.099	n.s.	
	Satisfaction	.099	n.s.	
Satisfaction	Promotion emotions	.215	n.s.	.401
	Prevention emotions	.256	.009	
	Delight	.137	n.s.	

^{a}Promotion emotions of excitement and cheerfulness (Higgins 1997, 2001).
^{b}Prevention emotions of confidence and security (Higgins 1997, 2001)

tarian benefits are primarily to avoid pain, whereas the goals of a promotion focus served by hedonic benefits are primarily to seek pleasure. Consistent with our predictions, we found that the type and the intensity of consumption emotions differ between hedonic and utilitarian benefits because they help fulfill different goals of seeking pleasure and avoiding pain. The results from the three studies prove a correspondence that links hedonic benefits with the promotion emotions of cheerfulness and excitement, leading to delight, and a correspondence that links utilitarian benefits with the prevention emotions of confidence and security, leading to satisfaction. To the best of our knowledge, this is the first study to demonstrate such a correspondence among product design benefits (hedonic versus utilitarian), postconsumption emotions (delight versus satisfaction), and loyalty.

Hedonic benefits and delight, and utilitarian benefits and satisfaction. Prior work on emotion theories has shown that delight and satisfaction are not the same feelings (Lazarus 1991; Oliver, Rust, and Varki 1997). The emotion of delight has been shown to accompany higher levels of arousal than satisfaction (Oliver 1997). As we discussed previously, the fulfillment of prevention goals leads to greater confidence and security—emotions low in arousal (Lazarus 1991; Roseman 1991). Conversely, the fulfillment of promotion goals leads to greater excitement and cheerfulness—emotions high in arousal (Lazarus 1991; Roseman 1991). Given the high-arousal levels of delight and the promotion emotions of excitement and cheerfulness, we predicted and found evidence that promotion emotions of excitement and cheerfulness are the primary antecedent feelings of delight. We also found that the promotion emotions fully mediate the relationship between hedonic benefits and delight (see Figure 2, Model 1). Similarly, consistent with our prediction, we found evidence that the low-arousal feelings of confidence and security from utilitarian benefits are the primary antecedent feelings of satisfaction. Furthermore, the prevention emotions fully mediate the relationship between the level of utilitarian benefits and satisfaction (see Figure 2, Model 2). In addition, we found that there was a significant difference between the correlation of arousal and promotion emotions and the correlation of arousal and prevention emotions. The intensity of arousal and the magnitude of its correlation were stronger with the promotion emotions than with the prevention emotions. This shows that the positive consumption of hedonic benefits leads to greater arousal than the positive consumption of utilitarian benefits. However, in the case of a negative consumption experience, greater arousal was associated with the consumption of utilitarian benefits (i.e., anger) and not hedonic benefits. This is consistent with the principle of precedence that states that fulfillment of prevention goals by utilitarian benefits is more important than fulfillment of promotion goals by hedonic benefits (Chitturi, Raghunathan, and Mahajan 2007).

Delight, promotion emotions, word of mouth, and repurchase intentions. This research proposed and tested a model of relationships involving promotion emotions, delight, word of mouth, and repurchase intentions (see Figure 2, Model 1). We also found that promotion emotions directly affect delight, word of mouth, and repurchase intentions. However, delight only partially mediates the relationship between promotion emotions and word of mouth. Furthermore, delight did not directly influence repurchase intentions in Model 1. The insight generated from the results of testing Model 1 (see Figure 1) is that even delight may not be an adequate measure of the total consumption experience from fulfillment of promotion goals. There are other aspects of product consumption experience captured by the promotion emotions of cheerfulness and excitement that directly affect word of mouth and repurchase intentions.

Satisfaction, prevention emotions, word of mouth, and repurchase intentions. We proposed and tested a second model of relationships involving prevention emotions, satisfaction, word of mouth, and repurchase intentions (see Figure 2, Model 2). The research found that prevention emotions directly affect satisfaction, word of mouth, and repurchase intentions. In addition, satisfaction partially mediates the relationship between prevention emotions and word of mouth. The relationship between prevention emotions and repurchase intentions is also partially mediated by satisfaction. The insight generated from the results of testing Model 2 (see Figure 1) is that satisfaction may not be an adequate measure of the total consumption experience from fulfillment of prevention goals. There are other aspects of product consumption experience captured by the prevention emotions of confidence and security that directly affect word of mouth and repurchase intentions.

Managerial Implications and Further Research

The research findings in this article could have significant implications for decision makers in product design and marketing organizations. Because of budget and time constraints, designers and managers are often compelled to choose among various attributes. If there is no budget or time constraints, perhaps the best solution is to maximize both hedonic and utilitarian dimensions of benefit. What is the limit for maximization? At what point do diminishing returns begin to affect the return on investment negatively? Are the diminishing returns significantly different for hedonic and utilitarian benefits? Does this fundamental difference lead to significant differences in the diffusion of innovations? These are potential topics for further research. These unknowns make the trade-off process between various attributes and benefits difficult and nonoptimal (Chitturi, Raghunathan, and Mahajan 2007). However, more often than not, product designers and managers are forced to trade off between selecting one attribute or another for various reasons. In such situations, we believe that the designers and marketing managers would be able to make better decisions if they considered the insights in this research.

As Keiningham and Vavra (2001, p. 176) state, "Creating delight for your customers first requires knowing and eliminating their points of pain, and then listening to their desires." Norman (2004) and Postrel (2003) propose the role of design in enhancing the emotional experience with the consumption of a product. Chaudhuri and Holbrook (2001) study the role of brand equity on brand performance in the context of the hedonic and utilitarian benefits of a product. Chaudhuri (2006) discusses a correspondence between merchandising value and utilitarian value and between differentiation value and hedonic value. Mahajan and Wind (2002) discuss an emotion-based product-positioning strategy. A discussion on how the insights developed in the current research could potentially bridge the gaps between managerial assertions and academic literature on design, delight, and loyalty follows.

Design strategy, delight, and satisfaction. The results from this research demonstrate a correspondence among utilitarian benefits, prevention emotions of confidence and security, and customer satisfaction. We find that customer delight is not directly influenced by either prevention emotions or satisfaction. Therefore, improving utilitarian benefits well beyond customers' needs and expectations might achieve high satisfaction but not necessarily delight. If designers and marketers want to delight customers, they should consider offering higher levels of hedonic benefits that fulfill promotion goals that lead to excitement and cheerfulness. Furthermore, we also find that satisfaction is directly influenced by prevention emotions but not by promotion emotions and delight.

What does it all mean for the design, development, and marketing of products? The process of product development and testing can be made much more benefits–experience–emotion centric than purely "attribute centric" by including the following steps: (1) Rate every attribute being considered for inclusion in the design specification on its ability and likelihood of fulfilling target customers' promotion and prevention goals; (2) use conjoint analysis to identify the most desirable combination of the promotion and prevention goals and their relative importance to maximize customer loyalty; (3) calibrate and match the attributes on the basis of their contribution rating on promotion and prevention dimensions from Step 1 with the ideal combination of promotion and prevention goals for the target customer segment from Step 2; and (4) iteratively calibrate concept testing, product use testing, and market testing with the ideal profile of promotion and prevention goals developed in Steps 1, 2, and 3 until an optimal product and marketing plan is achieved. We believe that the incorporation of these steps into the new product development process will improve the probability of success for a new product. As a potential topic of further research, it would be useful to know whether this proposition is beneficial and also whether this approach is more beneficial for radical versus incremental innovations.

Does delight improve loyalty? Although practitioners often assert that delight is essential for greater loyalty, the evidence in the literature is mixed (Rust and Oliver 2000). This research provides new insights into this issue. It shows that customer delight directly affects word of mouth and repurchase intentions. Therefore, delight improves two of the key measures of loyalty and must be pursued (Jacoby and Chestnut 1978). In addition, we find that promotion emotions that directly affect delight also directly affect word of mouth and repurchase intentions. This demonstrates the role of hedonics in improving loyalty by improving promotion emotions that also improve delight. Furthermore, analysis shows that both delight and satisfaction directly affect word of mouth and repurchase intentions. These findings suggest that there is more going on between the customer and the product than just satisfaction. We suggest that designers and marketers should pursue the goal of not just satisfying but also delighting the customer. In addition, improving the promotion emotions of cheerfulness and excitement and the prevention emotions of confidence and security will further improve loyalty. A related potential topic for further research could be to examine whether promotion emotions, prevention emotions, and delight improve loyalty for business-to-business products and services as well.

Is surprise necessary to delight customers? Chandler (1989) considers customer delight fundamentally different from customer satisfaction but claims that surprising (i.e., unanticipated) satisfaction is necessary for delight. In this research, we did not find evidence that surprise is necessary for delight, but we found evidence of arousal influencing delight. It is possible that arousal fully mediates the relationship between surprise and delight. If this conjecture is true, in the case of a completely utilitarian product, perhaps surprise is the only means to increase the arousal that is necessary for delight. Therefore, depending on the primary nature of the product (hedonic or utilitarian), marketers need to use either surprise or hedonic benefits as a potential means to delight. We also know that surprise is more difficult to sustain (Rust and Oliver 2000). Therefore, surprise-based delight could also be difficult to sustain. It would be useful to find out whether hedonics-based promotion emotions leading to delight are easier to sustain because they are not based on surprise. This is a potential topic for further research that has significant managerial implications for product design and marketing strategy.

Conclusion

On the basis of the results from our studies, we demonstrated a correspondence among hedonic benefits, delight, and loyalty. Similarly, we demonstrated a correspondence among utilitarian benefits, satisfaction, and loyalty. We also found evidence that though delight and satisfaction are positively correlated, they are significantly different in terms of their antecedent emotions—promotion and prevention emotions, respectively—and the consequences for customer loyalty, as measured by word of mouth and repurchase intentions. Finally, we showed that delighting customers improves word of mouth and repurchase intentions more than merely satisfying them.

APPENDIX A
Sample Items from Study Measures

Postconsumption Emotion Measures

"Based on the overall experience of using my car, I feel ..."

	Not at All						Extremely
Angry	1	2	3	4	5	6	7
Disappointed	1	2	3	4	5	6	7
Confident	1	2	3	4	5	6	7
Delighted	1	2	3	4	5	6	7

Word of Mouth

"Given the experience with your car, how likely are you to recommend your car to others?"

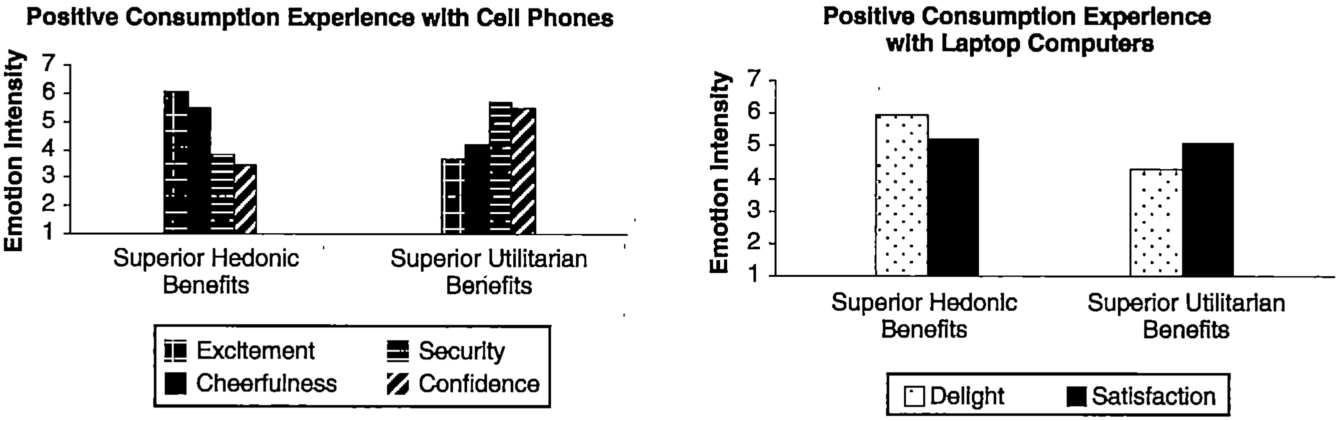
Not at All Likely						Extremely Likely
1	2	3	4	5	6	7

Repurchase Intentions

"Given the experience with your car, would you repurchase the same car, if you had to do it all over again?"

Not at All Likely						Extremely Likely
1	2	3	4	5	6	7

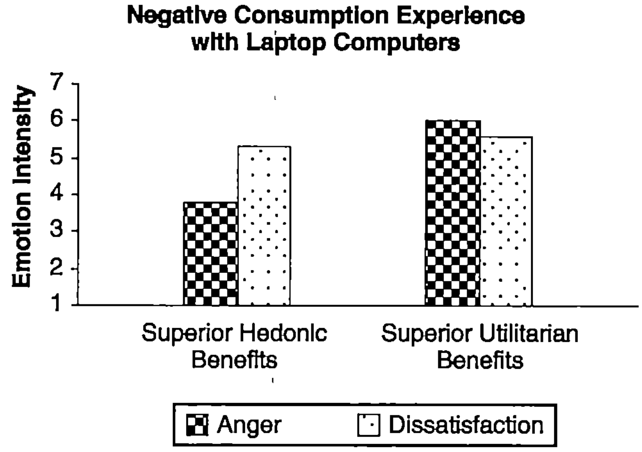
APPENDIX B
Intensity of Postconsumption Emotions for Products with Superior Hedonic Versus Superior Utilitarian Benefits



1: Study 1: Excitement–Cheerfulness Versus Confidence–Security

2: Study 2: Delight Versus Satisfaction

3: Study 2: Anger Versus Dissatisfaction



REFERENCES

Bagozzi, Richard P., Mahesh Gopinath, and Prashanth U. Nyer (1999), "The Role of Emotions in Marketing," *Journal of the Academy of Marketing Science*, 27 (2), 184–206.

Baron, Reuben M. and David A. Kenny (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations," *Journal of Personality and Social Psychology*, 51 (6), 1173–82.

Batra, Rajeev and O.T. Ahtola (1990), "Measuring the Hedonic and Utilitarian Sources of Consumer Attitudes," *Marketing Letters*, 2 (2), 159–70.

Bergkvist, Lars and John Rossiter (2007), "The Predictive Validity of the Multiple-Item Versus Single-Item Measures of the Same Constructs," *Journal of Marketing Research*, 44 (May), 175–84.

Berry, Christopher J. (1994), *The Idea of Luxury: A Conceptual and Historical Investigation*. New York: Cambridge University Press.

Chandler, Colby H. (1989), "Quality: Beyond Customer Satisfaction," *Quality Progress*, 22 (February), 30–32.

Chaudhuri, Arjun (2006), *Emotion and Reason in Consumer Behavior*. Oxford: Elsevier/Butterworth-Heinemann.

——— and Morris B. Holbrook (2001), "The Chain of Effects from Brand Trust and Brand Affect to Brand Performance: The Role of Brand Loyalty," *Journal of Marketing*, 65 (April), 81–93.

Chernev, Alexander (2004), "Goal-Attribute Compatibility in Consumer Choice," *Journal of Consumer Psychology*, 14 (1–2), 141–50.

Chitturi, Ravindra, Rajagopal Raghunathan, and Vijay Mahajan (2007), "Form Versus Function: How the Intensities of Specific Emotions Evoked in Functional Versus Hedonic Trade-Offs Mediate Product Preferences," *Journal of Marketing Research*, 44 (November), 702–714.

Crowley, Ayn E., Eric R. Spangenberg, and Kevin R. Hughes (1992), "Measuring the Hedonic and Utilitarian Dimensions of Attitudes Toward Product Category," *Marketing Letters*, 3 (3), 239–49.

Dhar, Ravi and Klaus Wertenbroch (2000), "Consumer Choice Between Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 37 (February), 60–71.

Frijda, Nico H. (1986), *The Emotions*. Cambridge, UK: Cambridge University Press.

——— (1987), "Emotion, Cognitive Structure, and Action Tendency," *Cognition and Emotion*, 1, 115–43.

Higgins, E. Tory (1997), "Beyond Pleasure and Pain," *American Psychologist*, 52 (12), 1280–1300.

——— (2001), "Promotion and Prevention Experiences: Relating Emotions to Nonemotional Motivational States," in *Handbook of Affect and Social Cognition*, Joseph P. Forgas, ed. London: Lawrence Erlbaum Associates, 186–211.

Holbrook, B. Morris (1999), *Consumer Value: A Framework for Analysis and Research*. New York: Routledge.

Hunt, H. Keith (1977), "CS/D: Overview and Future Research Directions," in *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*, H. Keith Hunt, ed. Cambridge, MA: Marketing Science Institute, 455–88.

Inman, Jeffrey J., James S. Dyer, and Jiamin Jia (1997), "A Generalized Utility Model of Disappointment and Regret Effects on Post-Choice Valuation," *Marketing Science*, 16 (2), 97–111.

Jacoby, Jacob and Robert W. Chestnut (1978), *Brand Loyalty: Measurement and Management*. New York: John Wiley & Sons.

Jones, Thomas O. and W. Earl Sasser Jr. (1995), "Why Satisfied Customers Defect," *Harvard Business Review*, 73 (November–December), 88–99.

Keiningham, Timothy and Terry Vavra (2001), *The Customer Delight Principle*. New York: McGraw-Hill.

Kivetz, Ran and Itamar Simonson (2002a), "Earning the Right to Indulge: Effort as a Determinant of Customer Preferences Toward Frequency Program Rewards," *Journal of Marketing Research*, 39 (May), 155–70.

——— and ——— (2002b), "Self-Control for the Righteous: Toward a Theory of Precommitment to Indulgence," *Journal of Consumer Research*, 29 (September), 199–217.

Larsen, Randy J. and Barbara L. Fredrickson (1999), "Measurement Issues in Emotion Research," in *Well-Being. The Foundation of Hedonic Psychology*, Daniel Kahneman, Ed Diener, and Norbert Schwarz, eds. New York: Russell Sage Foundation, 40–60.

Lazarus, Richard S. (1991), *Emotion and Adaptation*. New York: Oxford University Press.

Mahajan, Vijay and Yoram (Jerry) Wind (2002), "Got Emotional Product Positioning?" *Marketing Management*, 11 (May–June), 36–41.

Mano, Haim and Richard L. Oliver (1993), "Assessing the Dimensionality and Structure of Consumption Experience: Evaluation, Feeling, and Satisfaction," *Journal of Consumer Research*, 20 (December), 451–66.

Maslow, Abraham H. (1970), *Motivation and Personality*, 2d ed. New York: Harper and Row.

Mehrabian, Albert and James Russell (1974), *An Approach to Environmental Psychology*. Cambridge, MA: MIT Press.

Mittal, Vikas, William T. Ross Jr., and Patrick M. Baldasare (1998), "The Asymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions," *Journal of Marketing*, 62 (January), 33–47.

Norman, Donald A. (2004), *Emotional Design: Why We Love (or Hate) Everyday Things*. New York: Basic Books.

Okada, Erica M. (2005), "Justification Effects on Consumer Choice of Hedonic and Utilitarian Goods," *Journal of Marketing Research*, 42 (February), 43–53.

Oliver, Richard (1997), *Satisfaction: A Behavioral Perspective on the Consumer*. New York: McGraw-Hill.

———, Roland T. Rust, and Sajeev Varki (1997), "Customer Delight: Foundations, Findings, and Managerial Insight," *Journal of Retailing*, 73 (3), 311–36.

Ortony, Andrew, Gerald L. Clore, and Allan Collins (1988), *The Cognitive Structure of Emotions*. Cambridge, UK: Cambridge University Press.

Plutchik, Robert (1980), *Emotion: A Psychoevolutionary Synthesis*. New York: Harper and Row.

Postrel, Virginia (2003), *The Substance of Style*. New York: HarperCollins Publishers.

Reichheld, Frederick F. (1993), "Loyalty-Based Management," *Harvard Business Review*, 71 (March–April), 64–73.

Roseman, Ira J. (1991), "Appraisal Determinants of Discrete Emotions," *Cognition and Emotion*, 5 (3), 161–200.

Russell, James A. (1980), "A Circumplex Model of Affect," *Journal of Personality and Social Psychology*, 39 (6), 1161–78.

Rust, Roland T. and Richard L. Oliver (2000), "Should We Delight the Customer?" *Journal of the Academy of Marketing Science*, 28 (1), 86–94.

Strahilevitz, Michal and John G. Myers (1998), "Donations to Charity as Purchase Incentives: How Well They Work May Depend on What You Are Trying to Sell," *Journal of Consumer Research*, 24 (4), 434–46.

Voss, Kevin E., Eric R. Spangenberg, and Bianca Grohman (2003), "Measuring the Hedonic and Utilitarian Dimensions of Consumer Attitude," *Journal of Marketing Research*, 40 (August), 310–20.

Weber, Max (1998), *The Protestant Ethic and the Spirit of Capitalism*, 2d ed. Los Angeles: Roxbury.

Westbrook, Robert A. (1987), "Product/Consumption-Based Affective Responses and Postpurchase Processes," *Journal of Marketing Research*, 24 (August), 258–70.

——— and Richard L. Oliver (1991), "The Dimensionality of Consumption Emotion Patterns and Consumer Satisfaction," *Journal of Consumer Research*, 18 (June), 84–91.

Ulrich R. Orth & Keven Malkewitz

Holistic Package Design and Consumer Brand Impressions

This article develops empirically based guidelines to assist managers in selecting or modifying package designs for achieving desired consumer responses. Seven studies identify the key types of package designs, including the factors that differentiate those package designs, and determine how these holistic designs are related to consumer brand impressions. The selection of package designs can be simplified with the use of five holistic types: massive, contrasting, natural, delicate, and nondescript designs. Sincere brands should have natural package designs, exciting brands should have contrasting designs, competent brands should have delicate designs, sophisticated brands should have natural or delicate designs, and rugged brands should have contrasting or massive designs. The authors discuss the potential trade-offs among the impressions created by holistic design types and illustrate their findings with numerous real packages.

Keywords: package design, brand personality, gestalt, wine, perfume, visual competitor

Package design involves several considerations ranging from protecting package contents to articulating and communicating desired brand impressions. Because of this wide range, package design is a broad term spanning engineered functional attributes (e.g., ergonomics, durability, recyclability) and a package's visual attributes (Bloch 1995). Although we do not want to minimize the importance of other characteristics, the focus of this article is on design elements that create a package's visual appearance. This appearance is often an integral part of a brand's image, such as the hourglass shape of the Coca-Cola bottle and its logo in Spencerian type or the round-shouldered Bordeaux-style wine bottles that bear ornate labels and flourish typography.

Following Bloch (1995) and Gestalt psychologists (Koffka 1922; Wertheimer 1925), we define "package design" as the various elements chosen and blended into a holistic design to achieve a particular sensory effect. Designers choose design elements, decide how to mix them, and determine the desired level of congruity among them (Lawson 1983). For example, the design of the trademark Coca-Cola bottle includes the fluted surface with parallel vertical grooves, the hourglass shape, the greenish-hued glass, and the iconic Spencerian script spelling "Coca-Cola" on the face of the bottle. Similarly, the design of a Bordeaux-style wine bottle may include a slender bottle silhouette, an ornate label with elaborate typography, and a natural color scheme. In both cases, the overall effect of the package comes not from any individual element but rather from the gestalt of all elements working together as a holistic design.

A significant body of research attests to the importance that managers and scholars attach to package design (e.g., Bloch 1995; Garber, Burke, and Jones 2000; Hertenstein, Platt, and Veryzer 2005; Rettie and Bruwer 2000; Schoormans and Robben 1997). Taken together, the results suggest that package design is an extremely influential medium because of its pervasive impact on purchasers, its presence at the crucial moment when the purchase decision is made, and consumers' high level of involvement when they actively scan packages in their decision making. Of special interest for this research is package design's ability to assist in building strong brands by distinguishing offerings and by evoking favorable consumer impressions (Berkowitz 1987; Bloch 1995; Henderson et al. 2003; Schmitt and Simonson 1995). This ability to create and convey brand impressions has been studied by academics only recently (Henderson et al. 2003; Henderson, Giese, and Cote 2004; Underwood 2003). However, although extant research has demonstrated that package design can have a powerful impact on consumer response, it has provided only scant guidance to managers in selecting package design for creating desired responses.

Specifically, there are two major gaps in the knowledge. First, little research examining generalizable, holistic designs and their underlying factors exists. Identifying these archetypes, prototypes, or standards will substantially improve the understanding of package design, just as categorizing the vast number of individual traits into the "Big Five" traits facilitates the understanding of human personality. For example, numerous product categories include package designs that appear massive, natural, or delicate, at least relative to one another. Many consumer goods packages—from wine bottles and fragrances to cereals, detergents, and even products such as footwear or MP3 players—could all be differentiated according to these groupings. However, there appears to be a lack of meaning-

---

Ulrich R. Orth is Professor and Chair of A&F Marketing, Christian-Albrechts-Universität Kiel (e-mail: uorth@ae.uni-kiel.de). Keven Malkewitz is Assistant Professor of Marketing, College of Business, Oregon State University (e-mail: keven.malkewitz@bus.oregonstate.edu). This research was funded in part by Willamette Valley Vineyards. The authors gratefully acknowledge Andrea Marks, Carol Caughey, and Gerry Glasgow for their valuable comments and numerous professional designers for their input, especially Jay Thompson and Peter Moore. They thank Steffi Gal for her assistance in collecting the fragrances data, Cindy Lederer for providing access to the Oregon Consumer Panel, and the two anonymous *JM* reviewers for their insightful comments on previous versions of this article.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

ful guidelines to assist design stakeholders in creating holistic designs.

The second knowledge gap is related to the responses evoked by any given holistic design. A lack of insight into the relationships between key types of package designs and generic dimensions of consumer response leads to inefficiencies in achieving brand management objectives and leaves managers and designers in the dark in terms of what design to use to stimulate desired responses. To leverage design, companies need guidelines for achieving brand management goals, or at least a taxonomy of design types to help them consider the range of design options and resultant messages available to them. We are not suggesting that managers should design. Instead, by identifying prototypical package designs and their relationships to generic response dimensions, we allow managers to communicate better with designers and to provide designers with the structure they need to maximize their creativity.

To develop guidelines that address the previously outlined gaps, we begin by identifying design elements and factors that are important to and are used by designers in differentiating package designs. Next, we examine brand management literature to identify what types of responses managers desire. On the basis of these literature reviews and empirical assessments, we develop a series of studies to address two fundamental questions: (1) What are the key types of holistic package designs, including their differentiating design factors? and (2) How are these holistic package designs related to generalizable brand impressions?

Package Design Literature

There is both theoretical and empirical support for generalizable, prototypical, holistic package designs and their relationship to generalizable response dimensions. First, drawing from Gestalt psychology (Koffka 1922; Wertheimer 1925) and work on nonconscious preference formation (Kunst-Wilson and Zajonc 1980; Lewicki 1986), classification (Berlyne 1971), and type (Pepper 1949), current design processing theory suggests that specific design elements are perceived and organized into more complex components (Veryzer 1999). Second, several studies have directly explored fundamental factors that capture differences among key types of designs (Henderson and Cote 1998; Henderson et al. 2003; Henderson, Giese, and Cote 2004). Third, a significant amount of literature attests to package design's ability to generate a variety of impressions (e.g., Aaker 1991; Batra and Homer 2004; Batra, Lehmann, and Singh 1993; Keller 1993; Schmitt and Simonson 1995; Underwood 2003; Underwood and Klein 2002; Underwood and Ozanne 1998). Thus, we put forward the idea of generalizable holistic package designs, which are systematically related to generalizable brand impressions. Next, we review the supporting literature in greater detail.

Gestalt Psychology and the Part–Whole Distinction

Two features of the Gestalt psychology research on part versus whole perceptions are relevant to our work on holistic design types. First, the general idea of part–whole perceptual differences was one of the pioneering contributions of early Gestalt psychologists (Koffka 1922; Wertheimer 1925). For example, consider one colored dot versus a grouping of colored dots versus a whole painting in the classic pointillism style. A single painted dot might look ominous when seen alone, but it adds richness to a flower that is part of a still-life painting of a bouquet. Likewise, any one of the pure dots of color taken in isolation fails to convey the painting itself, but content and meaning emerge from the orchestration of numerous parts to construct a far richer perceptual whole. Second, Gestalt psychology also recognizes the importance of stimulus categorization, though more in terms of stimulus recognition. Consider figure–ground distinctions, according to which two designs might have similar features but differ in terms of which ones are made more prominent (the figure) and which are treated more as the background. A typical example is the so-called Rubin vase (named after Edgar Rubin, the Danish psychologist who studied such images), a vase–face, figure–ground reversal stimulus. On first sight, people might categorize the image as that of a white vase against a black background, whereas a closer look might reveal two human heads facing each other with contours of chins, lips, noses, and other facial features outlining the shape of the original vase.

Factors Underlying Holistic Package Designs

In addition to Gestalt psychology, recent design theory and empirical evidence suggest that consumers perceive "constitutive" elements (e.g., colors, textures, surfaces) and organize them into more complex components (Veryzer 1999). Such a hierarchical processing explanation draws from studies on nonconscious preference formation (e.g., Kunst-Wilson and Zajonc 1980; Lewicki 1986) and from findings that more abstract, multidimensional design characteristics are determined by more basic and measurable design elements (Geistfeld, Sproles, and Badenhop 1977). Design elements are combined into more complex (cognitive) components or factors of design, which are then aggregated during perception and convey particular characteristics (i.e., of a brand) to consumers. This interpretation is consistent with fluency-based attributions (Janiszewski and Meyvis 2001) and perceptions of beauty (e.g., Grammer and Thornhill 1994).

Furthermore, categorization, classification, and type theories suggest that generic dimensions exist during the progression that occurs as initial design perception moves to interpretation. Common to those theories is their reference to the recognition of a particular class of objects without the ability to identify all its details and peculiarities (Berlyne 1971) as consumers try to understand a stimulus by placing it within an existing category (Loken and Ward 1990). Thus, categorization is based on the perceived similarity between a given package and exemplars of various categories. A consumer first encountering a package may initially categorize the design as belonging to a particular category, but as perception becomes more complete and the object is more fully understood, the person moves past the generic design or category stage. Similarly, generic holistic package design may be understood in terms of Pepper's (1949) "type" concept. Type refers to an association of a certain set of traits in certain relationships, such that they are recognizable as a whole. Every generic design is defined by the association of a certain set of design factors and elements in certain relationships.

In addition to theoretical considerations, several empirical studies support the emergent part-to-whole perceptual properties of the design elements → generalizable factors → holistic designs sequence. For example, Henderson and Cote (1998) generate insights that assist managers in selecting or modifying logo designs to achieve corporate image goals. Categorizing almost 200 logos on the basis of 22 design elements, their analysis yielded seven underlying factors. Three of the seven factors were significant. The first factor, "natural," combines lower-level characteristics, such as representative and organic, and reflects the degree to which constitutive design elements depict commonly experienced objects. For example, a logo consisting of an image of a leaf would be judged to be more natural than one consisting of a building. The second factor, "harmony," combines symmetry and balance. A logo that uses design elements that are symmetric around both a horizontal and a vertical axis would be more harmonious than one that does not. The third factor, "elaborate," is a combination of design element complexity, activity, and depth. This factor captures the concept of design richness and the ability of the design elements to capture the essence of something. Examining logos in Asia, Henderson and colleagues (2003) find eight design characteristics captured by three underlying design dimensions. Labeled "elaborate," "harmony," and "natural," these factors are essentially identical to those of Henderson and Cote (1998). Despite the different cultural contexts of the studies, only minor differences were observed between the factor compositions.

Extending the logo studies, Henderson, Giese, and Cote (2004) extract six factors from an extensive set of typefaces. Three "universal" factors were judged to be applicable to all design, and three factors were specific to typeface design. Their general factors resemble Henderson and Cote's (1998) findings, though they use different, practitioner-specified design elements. The first factor, "elaborateness," includes positive loadings for ornate, depth, distinctive, and meaningful and negative loadings for readable and common. The second factor, "harmony," includes balance, smoothness, symmetry, and uniformity. The third factor, "naturalness," includes positive loadings for active, curved, organic, and slant and a negative loading for typed.

Other studies have more narrowly focused on specific design elements, including color (Garber, Hyatt, and Starr 2000), imagery (Underwood and Klein 2002), shape (Folkes and Matta 2004; Yang and Raghubir 2005), size (Folkes and Matta 2004; Wansink 1996), proportions (Raghubir and Greenleaf 2006), unity, and prototypicality (Veryzer and Hutchinson 1998), but they do not establish links to generic design factors. In summary, the literature reviewed reveals that package design should be studied from a holistic perspective—namely, in terms of higher-order design factors that differentiate generalizable holistic package designs. Translating basic design elements into higher-order factors and more understandable holistic designs, including the impressions they evoke with consumers, is critical in helping firms better understand package design and consumers' responses to it.

Package Design and Consumer Impressions

The design of a package elicits various responses from consumers (for a review, see Bloch 1995). Different from previous work that addresses the aesthetic aspects of design (e.g., Hirschman 1986; Holbrook 1986; Veryzer and Hutchinson 1998) or focuses on global cognitive evaluations and their effects on consumer behavior (e.g., Folkes and Matta 2004; Garber, Burke, and Jones 2000; Rahgubir and Krishna 1993; Schoormans and Robben 1997), our research focuses on generalizable consumer brand impressions inferred from package design.

The literature has documented numerous impressions generated by packages, and design has been repeatedly emphasized as an instrument for conveying meaning and generating brand impressions (Aaker 1991; Batra and Homer 2004; Schmitt and Simonson 1995). Depending on the research context, these impressions are related to the terminology of designers (Forty 1992), to brand management objectives (Underwood 2003), or to consumer benefits (Batra and Homer 2004; Batra, Lehmann, and Singh 1993). Many researchers emphasize that a particular design should evoke the same intended meaning across people (e.g., Borja de Mozota 2003; Henderson and Cote 1998; Henderson et al. 2003; Henderson, Giese, and Cote 2004; Rettie and Bruwer 2000; Schoormans and Robben 1997; Underwood 2003; Underwood and Klein 2002; Underwood and Ozanne 1998). This perspective is consistent with current marketing thought that brand communications should convey a single, clear message that is difficult to misinterpret (Keller 1993; Underwood 2003).

Brand personality has emerged as a key tool to capture and categorize facets of brands systematically in terms of generalizable impressions responses (Aaker 1997, 1999; Aaker, Benet-Martinez, and Garolera 2001; Aaker, Fournier, and Adam 2004). Consumer researchers have given a considerable amount of attention to the concept, focusing on how brand personality enables consumers to express their actual self (Belk 1988; Sirgy 1982), ideal self (Elliott and Wattanasuwan 1998), or specific dimensions of the self through brand acquisition and use (Fournier 1998; McCracken 1986). In her seminal article, Aaker (1997) conceptualizes brand personality and explores the meaning of five basic dimensions: "sincerity," "excitement," "competence," "sophistication," and "ruggedness." Subsequent studies confirm the number and nature of these dimensions (Sweeney and Brandon 2006). Managers also view brand personality as both a means of differentiating a brand (Vernadakis 2000) and a common denominator for marketing a brand across cultures (Aaker, Benet-Martinez, and Garolera 2001; Sung and Tinkham 2005). Insights into the processes of personality inferences about brands further emphasize the role of design for brand personality impressions (Johar, Sengupta, and Aaker 2005). Package designs are especially critical for impressions formed for new brands (Underwood and Klein 2002). However, although the brand personality concept appears well suited to capture generalizable response dimensions, there are no guidelines on how a firm can develop holistic package designs for achieving desired brand personality impressions.

Research Goal

Brand package modifications typically set out from the impressions that management wants to communicate through the design. Designers are then charged with developing packages that are believed to evoke the desired impressions. However, the question remains whether designs developed through executive experience and intuition effectively generate desired responses. This research empirically investigates how firms can develop package designs for achieving desired responses.

Method

We apply methods previously used and accepted in marketing research on design (Henderson and Cote 1998; Henderson et al. 2003; Henderson, Giese, and Cote 2004). As do prior studies, we employed two-dimensional stimuli from a suitable product category for identifying design characteristics useful both within and outside the stimulus category. This method is not used just in design research; it is analogous to advertising studies using one exemplary stimulus to examine consumer responses and to derive more generalizable conclusions. The methodology involves (1) selecting a product category appropriate for the research question, (2) determining representative stimulus characteristics, (3) selecting stimuli for the research, (4) obtaining expert ratings of them, and (5) assessing consumer responses to the stimuli.

Product Category Selection

Four major concerns drive the selection of an appropriate example product. First, to ensure generalizable results, a branded product is required with package design elements commonly found in a wide variety of consumer goods. Second, the product should provide brand names that are unfamiliar to the sample population to avoid confounds with different levels of brand familiarity (Underwood and Klein 2002). Third, a product is needed for which package design has a measurable impact on consumer impressions and purchase decisions. Fourth, and perhaps most important, there needs to be large variance in package designs in the selected category, permitting a meaningful assembly of actual stimuli.

On the basis of these criteria, we selected stimuli from the wine category for our research. Wine bottles contain a wide variety of design elements commonly found in other consumer goods, there is an opportunity to use brands largely unfamiliar to research participants, there is at least anecdotal evidence that wine package design influences consumer impressions, and the large variance in wine packages and brand personalities allows for a meaningful assembly of actual stimuli designs.

Representative Package Design Elements

Because our research aims to identify prototypical holistic package designs that are generalizable for many types of consumer goods, we obtained an initial list of design elements from previously reviewed literature. We expanded this list using trade and academic publications on wine brand management and package design (e.g., Caldewey and House 2003; Caputo 2005; Mackay 2005). Next, nine professional designers from multiple firms and with experience in a broad range of design tasks, including wine package design, listed primary differentiating elements of package design in general and of wine packages in particular. We integrated their feedback with the initial list, which resulted in a final list of 62 design elements. These elements were the basis for identifying holistic package designs and underlying factors.

Stimuli

To prepare the selection of stimuli, we asked each of the designers involved in the selection of elements to assemble a list of wine packages representative of the variance in the final design elements list (e.g., one package design with heavy typography, a second with light typography, a third with a small label, a fourth with a large label). Because consumers rely more on package design elements for inferring brand characteristics of unfamiliar brands than of familiar brands (Underwood and Klein 2002), we reduced the expansive pool of wine packages to brands with little familiarity. We used additional input from a second set of professionals in the wine and design industries to select a total of 160 wine bottles that represented the full range of the 62 design elements. These package designs were purchased from various retail sources.

Stimuli for the subsequent data collection were produced by taking high-resolution digital images of the packages. Images were taken in a single session in a professional photo studio before a matte monochromatic background. To preserve size and proportion differences, the tallest bottle served as a calibration standard. All images were framed and taken with the object and camera in the same location; there was no variation in illumination, timing, or any other variable.

Expert Ratings of Package Design Elements

A third set of professional designers (N = 125) associated with different firms each rated a subset of no more than 12 of the 160 stimuli on half of the 62 design elements, using a seven-point semantic differential scale. This task took approximately 20 minutes. Digital images were posted on a Web site accessible only through a confidential login and password. Software was developed to assign 12 stimuli randomly to each designer. Each package was presented individually on the left-hand side of the screen, and the respondents scrolled through the semantic differential scales on the right. Participants controlled the amount of time for viewing and responding. To ensure similarity in appearance across different hardware settings, images were optimized for Internet Explorer. In all, the professionals provided

38,750 individual ratings (approximately 9 designers per stimulus). Descriptive results appear in Table 1.

Relevant Consumer Impressions

Our literature review highlighted the importance of brand personality for brand differentiation and assessment of meaning (Aaker 1999; Fournier 1998; Johar, Sengupta, and Aaker 2005; Keller 2003; Sung and Tinkham 2005; Sweeney and Brandon 2006). To ascertain whether the response dimensions were detailed but generalizable, we included ten more impression items from managerially oriented literature. These impressions were "high quality," "corporate," "everyday," "cheap," "feminine," "evokes happy memories," "healthy," "stylish," "will impress my friends," and "value for money" (Orth 2005; Orth, Wolf, and Dodds 2005). For example, in many product categories, consumers may prefer small or family-owned companies over "corporate" brands. Others distinguish between brands for everyday consumption and those for special occasions (Quester and Smart 1998). Consumers also buy brands that remind them of good times and to relive memories of happy vacations or a carefree childhood. To confirm the relevance of the selected impressions, we asked professionals in the design industry whether the impressions were meaningful and whether they could design packages for creating these impressions. The results confirmed our selection of design elements and brand impressions as inclusive, actionable, and generalizable.

Consumer Response to Package Designs

We used a consumer survey to measure impressions evoked by package design. To present stimuli uniformly and to expose consumers to the exact stimuli evaluated by designers, we replicated the Web-based computer task. We modified the software algorithm to randomly present 10 (rather than 12) stimuli accompanied by the scroll-down impressions items. We used a consumer panel maintained by a research university in the Pacific Northwest, and the sample was representative for that area. To increase motivation and involvement with the study tasks, respondents received $10 gift certificates. The overall response was 67,000 individual ratings from 268 respondents. On average, respondents were 41.4 years of age, 57% were female, and only a small fraction was students.

Analysis and Results

Data analysis followed procedures established in research on design (Henderson and Cote 1998; Henderson et al. 2003; Henderson, Giese, and Cote 2004), experimental aesthetics (Seifert 1992), and language processing (Carpenter and Miyake 1995). Unique to this approach, analyses are conducted at the stimulus level rather than at the individual level. To obtain a score for each stimulus on a specific variable, we averaged individual ratings of a package (i.e., designer ratings and consumer responses). We conducted all remaining analyses using these averaged scores. Thus, the unit of analysis is the package design, and the sample size for each analysis is the number of different designs (i.e., the 160 stimuli). The mean brand familiarity score across all stimuli indicated that consumers were not familiar with the brands (M = 1.74, SD = .73; scale ranged from 1 = "not at all familiar" to 7 = "very much familiar").

Prototypical Holistic Package Designs

The first analytical step identified holistic package designs on the basis of similarities of the stimuli's 62 design elements. Because the combination of design elements drives holistic package design rather than any single element, we employed cluster analysis for this procedure (Henderson, Giese, and Cote 2004). We determined the number of clusters by examining the average distance between clusters and comparing this with the within-cluster distances. In addition, we avoided creating clusters with too few package designs. Five clusters appeared to best describe the data. We acknowledge that other design types or classifications exist, but we focus on more common types, consistent with the study's goal of identifying prototypical package designs and relating them to generalizable brand impressions.

In the second step, we performed exploratory factor analysis (EFA) with Varimax rotation for identifying design factors that are instrumental in differentiating the five holistic design clusters. On the basis of theory and previous research (Henderson and Cote 1998; Henderson et al. 2003; Henderson, Giese, and Cote 2004), we split design elements into two groups, one that included elements consistent with previously identified universal design factors (i.e., elaborate, harmony, and natural) and one that included all others. To extend previous research and advance design research across stimuli, we analyzed the two groups of design elements separately to ensure that we identified and linked generalizable dimensions to responses.

The factor analysis of the universal design factors group revealed three dimensions that explained 78.9% of the variance (see Table 2). The first factor, natural, included color scheme naturalness, organic versus geometric typography, bottle glass color naturalness, and images of nature, including landscapes, plants, and vines. The second factor, harmony, included image contrast, typography uniformity, color scheme harmony, image resolution, and brand name–logo contrast. The third factor, elaborate, included label degree of structure, amount of detail (image and label), typography elaboration, quantity of text, single versus multiple labels, label elaboration, and rounded versus angular label shape. The resultant design factors are similar to Henderson, Giese, and Cote's (2004) findings, though we used package design elements that industry professionals specified. A subsequent confirmatory factor analysis (CFA) confirmed the appropriateness of the factor structure (comparative fit index [CFI] = .891).

Factor analysis of the second group of design elements uncovered five additional dimensions that explained 69.2% of the variance (see Table 2). "Size" comprised label size, brand name/logo size, image size, label coverage, and horizontal versus vertical orientation of the label. "Symmetry" comprised label and bottle symmetry, as well as bottle silhouette, label material, and bottle color. Although it might appear that symmetry and harmony tap the same dimension, they were correlated at only .158. "Compressed" comprised

TABLE 1
Package Design Elements and Their Evaluation by Professionals

Design Element (Semantic Anchors)	M	SD
Adornment/charm presence (weak versus strong)	3.56	1.23
Awards information presence (weak versus strong)	1.88	1.10
Capsule presence (weak versus strong)	3.74	1.02
Closure presence (weak versus strong)	3.34	.85
Bottle		
–body (short versus long)	4.41	1.19
–girth (small versus large)	4.19	.90
–glass color (common versus unique)	3.46	.98
–glass color intensity (weak versus strong)	4.61	1.18
–glass color naturalness (low versus high)	3.84	.66
–glass finish (matte versus glossy)	4.95	.93
–glass transparency (low versus high)	3.16	1.23
–head/lip (not at all versus very much present)	4.13	.96
–heft/gravitas (light versus heavy)	4.33	.74
–neck (short versus long)	3.95	1.13
–relief/molding (not at all versus very much present)	3 17	1.16
–shoulders (short versus long)	3.87	1.33
–sides (straight versus rounded)	3.53	1.12
–silhouette (common versus unique)	3.54	1.29
–symmetry (asymmetrical versus symmetrical)	5.19	1.09
Brand name/logo		
–contrast (low versus high)	4.41	1.07
–location (low versus high)	3.64	1.23
–size (small versus large)	4.17	.99
Color scheme		
–harmony (inharmonious versus harmonious	4.37	1.28
–naturalness (not at all versus very natural)	4 24	1.11
–number of colors (monochromatic versus colorful)	3.45	1.27
–strength (weak versus strong)	4.33	.88
Image(s)		
–amount of detail (small versus large)	3.88	1.17
–execution (traditional versus contemporary)	4.03	1.67
–image–background contrast (low versus high)	4.05	1.04
–relation to animal(s) (weak versus strong)	2.02	1.72
–relation to building(s) (weak versus strong)	2.14	1.41
–relation to human being(s) (weak versus strong)	2.78	1.78
–relation to landscape(s) (weak versus strong)	2.64	1.65
–relation to nature (weak versus strong)	3.36	1.82
–relation to plant(s) (weak versus strong)	2.67	1.68
–resolution (low versus high)	3.92	.86
–size (small versus large)	4.09	1.15
–relation to wine (weak versus strong)	3.50	1.76
Image versus typography (image only versus text only)	4.23	1.09
Label(s)		
–amount of detail (small versus large)	4.08	1.28
–degree structure (low versus high)	4.33	.78
–elaboration (plain versus ornate)	4.21	1.34
–fragmentation (single label versus multiple labels)	2.89	1.71
–quantity text (small versus large)	3.45	1.10
–orientation (horizontal versus vertical)	3.93	1.38
–shape (rounded versus angular)	4.83	1.30
–surface area (little versus extensive coverage)	4.26	1.25
–symmetry (asymmetrical versus symmetrical)	4.66	1.12
Neck label: presence (weak versus strong)	3.18	1.53
Origin information: presence (weak versus strong)	4.05	1.39
Primary label		
–finish (matte versus glossy)	3.08	1.16
–material (common versus unique)	3.87	1.19
–position (low versus high)	3.43	1.07
–size (small versus large)	4.28	.98
–surface texture (smooth versus textured)	3.12	1.08
Typography		
–compression (condensed versus extended)	3.87	.84
–elaboration (plain versus ornate)	3.96	1.12
–flourishness (not flourish versus very flourishy)	3.69	1.08
–harmony (not uniform versus uniform)	4.38	.83
–naturalness (organic versus geometric)	3.89	1.06
–weight (light versus heavy)	3.85	.81
Varietal name. presence (weak versus strong)	4.33	1.24

TABLE 2
EFA of the Design Elements

Factors	Elements	Semantic Anchors	Loadings
Natural	Color scheme	Natural–not natural	.885
	Typography	Organic–geometric	–.857
	Bottle glass color	Natural–not natural	–.783
	Images of nature	Weak–strong	.765
	Images of landscapes	Weak–strong	.670
	Images of plants	Weak–strong	.647
	Images of wine	Weak–strong	.539
Harmony	Image contrast	Low–high	–.851
	Typography	Not uniform–uniform	.781
	Color scheme	Inharmonious–harmonious	.686
	Image resolution	Low–high	.653
	Brand name/logo contrast	Low–high	–.618
Elaborate	Label degree of structure	Low–high	.811
	Image amount of detail	Small–large	.799
	Typography	Plain–ornate	.703
	Label amount of detail	Small–large	.702
	Label quantity text	Small–large	.684
	Number of labels	Single–multiple	.663
	Label	Plain–ornate	.662
	Label shape	Rounded–angular	.652
Size	Label size	Small–large	.893
	Brand name/logo size	Small–large	.891
	Image size	Small–large	.769
	Label coverage	Little–extensive	.764
	Label orientation	Horizontal–vertical	.691
Symmetry	Label symmetry	Asymmetrical–symmetrical	.821
	Bottle symmetry	Asymmetrical–symmetrical	.754
	Bottle silhouette	Common–unique	–.746
	Label material	Common–unique	–.661
	Bottle color	Common–unique	–.603
Compressed	Label position	Low–high	.821
	Brand name/logo location	Low–high	.776
	Bottle neck	Short–long	–.712
	Typography	Condensed–extended	.631
Flourish	Bottle lip	Not at all–very much present	.837
	Typography flourishness	Not at all–very much flourish	.800
	Bottle relief	Not at all–very much present	.797
	Awards presence	Weak–strong	.725
	Adornment presence	Weak–strong	.662
	Neck label presence	Weak–strong	.658
	Closure presence	Weak–strong	.653
	Capsule presence	Weak–strong	.644
	Varietal name presence	Weak–strong	.569
Weight	Typography weight	Light–heavy	.746
	Bottle heft	Light–heavy	.745
	Bottle girth	Small–large	.663
	Bottle sides	Straight–rounded	.631
	Bottle body	Short–long	–.613

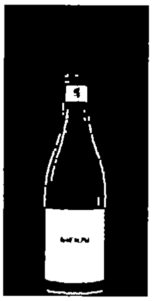
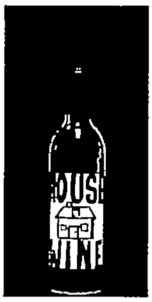
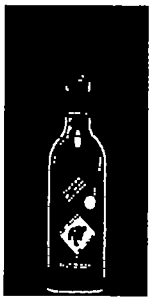
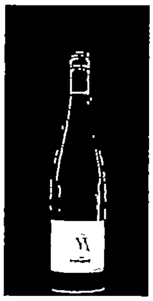
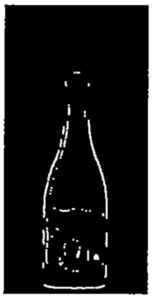
Notes: Factor structures resulting from initial EFAs with Varimax rotation were confirmed through subsequent CFA.

label position, brand name/logo location, bottle neck length, and condensed versus extended typography. It correlated with symmetry at only .181. "Flourish" comprised the presence of lips, awards, closures, capsules, varietal names, adornments, reliefs, and neck labels, as well as typography flourishness. The appearance that flourish and elaborate would tap the same dimension is contradicted by a weak correlation at only .171. The final factor, "weight," comprised typography weight, bottle heft, girth, sides, and bottle body. These dimensions are also similar to design factors and intermediate design characteristics established in previous studies. In particular, the factors flourish and weight (Henderson and Cote 1998; Henderson et al. 2003; Henderson, Giese, and Cote 2004), size (Folkes and Matta 2004; Wansink 1996), and compressed (Henderson and Cote 1998; Raghubir and Greenleaf 2006) appear in prior research. However, our results do not replicate the shared properties of symmetry and elaborate. Again, a CFA confirmed the appropriateness of the factor structure (CFI = .803). Because orthogonal factor scores could be created, we used the EFA results for subsequent examinations of the relationships between design and response.

To determine which design factors significantly differentiated clusters, we used analysis of variance. In addition, we performed t-tests to determine what cluster-specific factor scores were significantly smaller or greater than the mean score across all clusters. Table 3 presents the results of these analyses, including example packages.

The first cluster, labeled "massive," comprises 25 of the 160 package designs. Design factors that differentiate this holistic package design from others include above-average size, weight, and compressions and below-average natural, elaborate, and flourish. These factors capture broader groupings of design elements that characterize the holistic design type as one that can be best described as massive or robust. Exemplary package designs include House Wine, Meditrina, and Wine by Joe. The second cluster, "contrasting," is the smallest group in the sample, comprising 17 package designs. Differentiating factors include below-average harmony, natural, flourish, and compressed. These factors characterize the holistic package design as one that stands in stark contrast to the harmony factor identified in previous research. Exemplary contrasting designs include the "critter labels" Yellow Tail and Bear Crossing and other irregular designs, such as Ferngrove. The third cluster, "natural," is the largest group in the sample, comprising 45

TABLE 3
Holistic Package Designs and Differentiating Factors

	Holistic Design						
Design Factor	**Massive (N = 25)**	**Contrasting (N = 17)**	**Natural (N = 45)**	**Delicate (N = 32)**	**Nondescript (N = 33)**	**Sample Mean**	***p***
Natural	2.17^{-}	2.58^{-}	4.84^{+}	2.14^{-}	2.32^{-}	3.04	.001
Harmony	4.19	3.84^{-}	5.03^{+}	3.67^{-}	4.22	4.16	.001
Elaborate	3.10^{-}	4.03	4.30^{+}	4.60^{+}	3.43^{-}	3.83	.001
Size	4.84^{+}	3.99	4.03	3.09^{-}	3.89	3.92	.001
Symmetry	3.18	3.35	3.91^{+}	3.47	3.72^{+}	3.49	.003
Compressed	4.44^{+}	4.46^{+}	3.82	3.22^{-}	4.01	3.91	.001
Flourish	2.61^{-}	2.90^{-}	3.89^{+}	3.91^{+}	3.60	3.51	.001
Weight	4.43^{+}	4 13	3.88	3 78^{-}	3.89	3.98	.006
Examples							
							
							

Notes. The superscript "+" indicates that the cluster mean scores were significantly greater ($p <$ 01) than the pooled mean, and the superscript "–" indicates that the cluster means were significantly smaller than the pooled mean package designs. Differentiation is achieved through above-average natural, harmony, elaborate, symmetry, and flourish. Overall, this holistic design can be described as natural, representative, or archetypical. Typical examples include "old-world" package designs, such as Chateau Lagarenne, but also packages from "new-world" companies, such as Chateau Ste. Michelle and The Pines. The fourth cluster, "delicate," comprises 32 package designs. Factors differentiating this holistic cluster from others include above-average elaborate and flourish and below-average natural, harmony, size, weight, and compressed. Overall, this prototypical holistic design can be described as muted, sleek, and delicate. Typical package designs in this group include Travaglini, Mystic Wines, and Prinz von Hessen. The final cluster, "nondescript," comprises 33 package designs. Prominent within this holistic package design is the scarcity of outstanding design characteristics. Below-average natural and elaborate and above-average symmetry differentiate this design type from others. The main difference from the "delicate" cluster lies in the combination of design elements and factors producing little differentiation rather than distinctly delicate images. Package designs within this group can be described as simple, clean, discrete, and transparent. Typical package designs include Bierzo, Fusee, and Saint M.

Brand Impressions Responses

Brand personality data were analyzed through a CFA to capture generalizable response dimensions (Aaker 1997). The analysis produced five factors that corresponded to Aaker's sincerity, excitement, competence, sophistication, and ruggedness dimensions. The model fits the data reasonably well (CFI = .93), and all factor loadings were acceptably high (>.70). On the basis of these results, we used five brand personality dimensions to capture generalizable impression responses. Table 4 shows descriptive results and example designs. For example, Temptress received the lowest score on sincerity, and Sawtooth scored the highest on this dimension. Similarly, Bierzo generated the lowest on excitement, and Siskiyou's Sideways generated the highest.

It might be argued that the information collected on ten additional brand impressions was more wine specific; however, these responses are desirable for many products and assist in differentiating not only food and beverage packages but also products and packages such as fragrances, watches, and MP3 players. Because the goal of this research is to show that prototypical holistic package designs are related to generalizable response dimensions, we examined the relationships between brand personality dimensions and the managerially oriented brand impressions.

First, an EFA generated four managerially based impressions that loaded onto two factors, explaining 70.1% of the variance. Labeled "prestige," the first factor comprised stylish and impressive. The second factor, "inexpensive," comprised everyday and cheap. We averaged item scores for these factors, which resulted in a reduction of the original ten managerially based impressions to eight.

Second, we performed five stepwise hierarchical regression analyses with the managerially based impressions as predictor variables, one each for the five brand personality dimensions. The results indicate hierarchical relationships between brand personality dimensions and managerially based impressions. They also indicate that much of the information captured by the managerially based impressions is represented in the brand personality dimensions. Overall, percentages of variance explained ranged from .31 (ruggedness) to .88 (sophistication). For example, ruggedness (F = 23.22) is a function of feminine (–.60), evokes happy memories (.41), and inexpensive (.26). Sophistication (F = 185.12) is a function of quality (.42), prestige (.38), feminine (.18), inexpensive (–.16), and healthy (.10). Although these findings indicate that the five brand personality dimensions may sufficiently represent the desired generalizable response dimensions, we decided to include the eight managerially based impressions in further analyses because (1) retaining additional impressions that are based on managerial language and thinking provides more detailed insight into responses generated through package designs and (2) it preserves the richness of the data, considering that not all variance in the data is explained by the hierarchical relationships between brand personality dimensions and managerially based impressions. In addition, responses such as high quality, corporate, healthy, and value for money are desirable across many product categories and thus provide more generalizable insights.

Relationships Between Holistic Package Designs and Generalizable Response Dimensions

To show that generic holistic designs are associated with generalizable responses, we performed an analysis of variance. We found differences in consumer responses to the holistic package designs for both brand personality and managerially based impressions (Table 5).

Massive package designs are associated with impressions of excitement and score low on competence and sophistication. Sincerity and ruggedness are only average. Consumers further perceive brands with massive packages as low in quality, inexpensive, less healthy, and not classy. Contrasting package designs score high on excitement, low on competence and sophistication, and average on sincerity. In contrast to massive designs, contrasting packages lead to stronger impressions of ruggedness. Other differentiating impressions include low quality, not feminine, and inexpensive. Natural designs generate impressions of sincerity, competence, and sophistication; they score low on excitement and average on ruggedness. Natural designs are perceived as high quality, feminine, healthy, and expensive but a good value for money, and they evoke happy memories. Delicate package designs score high on competence and sophistication, whereas sincerity, excitement, and ruggedness responses are only average. Delicate designs further generate impressions of high quality, classy, and expensive. Finally, nondescript designs score low on sincerity, excitement, and ruggedness and average on competence and sophistication. These designs further generate impressions of corporate and little value for money and do not evoke happy memories.

Discussion of Findings

In previous research stages, we identified prototypical holistic package designs and their relationships to general-

TABLE 4
CFA Results for Brand Personality Dimensions

Factor	Item	Loading	M (SD)	Low	Example	High	Example
Sincerity	Down-to-earth Honest Wholesome Cheerful	.82 .89 .82 .74	3.01 (.89)	Temptress (1.88)		Sawtooth (4.12)	
Excitement	Daring Spirited Imaginative Up-to-date	.89 .91 .89 .72	3.04 (1.04)	Bierzo (1.82)		Siskiyou Sideways (4.42)	
Competence	Reliable Intelligent Successful	.71 .89 .92	3.16 (.97)	Punk Floyd (1.42)		Chateau Lagarenne (3.98)	
Sophistication	Upper class Charming	.85 .88	3.01 (1.03)	House Wine (1.97)		Griffin Creek (3.94)	
Ruggedness	Outdoorsy Tough	.86 .89	2.34 (1.06)	Bishop Creek (1.38)		Prosperity Red (4.11)	

izable response dimensions. Using wine as a carefully selected example product, we determined a universe of package design elements, selected representative designs as stimuli, collected professionals' ratings of package design elements, identified generalizable response dimensions, and assessed consumers' design-evoked impressions. Cluster

TABLE 5

Relationships Between Prototypical Holistic Package Designs and Generalizable Response Dimensions

	Holistic Design						
Response Dimension	**Massive (N = 25)**	**Contrasting (N = 17)**	**Natural (N = 45)**	**Delicate (N = 32)**	**Nondescript (N = 33)**	**Sample Mean**	***p***
Sincerity	2.96	3.05	**3.18^{a}**	2.94	**2.86^{a}**	3.02	.006
Excitement	**3.21^{abc}**	**3.56^{adef}**	**2.83^{bdg}**	3.16^{egh}	**2.82^{cfh}**	3.04	.001
Competence	**2.77^{abc}**	**2.89^{def}**	**3.35^{ad}**	**3.33^{be}**	3.17^{cf}	3.16	.001
Sophistication	**2.69^{abc}**	**2.72^{de}**	**3.22^{adf}**	**3.17^{be}**	2.96^{cf}	3.01	.001
Ruggedness	2.28^{a}	**2.83^{abcd}**	2.31^{b}	2.30^{c}	**2.21^{d}**	2.34	.001
MANOVA							.001
High quality	**3.70^{abcd}**	**4.07^{aefg}**	**4.86^{beh}**	**4.77^{cfi}**	4.44^{dghi}	4.47	.001
Corporate	**3.20^{abcd}**	3.78^{aef}	4.15^{be}	4.10^{c}	**4.23^{df}**	3.96	.001
Inexpensive	**4.32^{abc}**	**4.04^{def}**	**3.52^{ad}**	3.43^{be}	3.65^{cf}	3.72	.001
Feminine	3.36	**2.69^{a}**	**3.47^{a}**	3.10	3.10	3.21	.020
Happy memories	3.47	3.52	**3.80^{a}**	3.52	**3.27^{a}**	3.54	.005
Healthy	**3.67^{a}**	3.74	**4.11^{ab}**	3.99	3.78^{b}	3.90	.003
Prestige	**3.81^{abcd}**	4.11^{ae}	4.34^{b}	4.45^{cef}	4.12^{df}	4.21	.001
Value for money	3.88	4.12	**4.15^{a}**	4.05	**3.87^{a}**	4.02	.028
MANOVA							.001

Notes: Values in bold distinguish the cluster in terms of emphasizing that dimension score to be significantly ($p < .05$) greater or smaller than the sample mean. Within each dimension, identical superscripts identify pairs of designs that score significantly ($p < .05$) different on this factor on the basis of a Scheffé post hoc test. MANOVA = multivariate analysis of variance.

analysis generated five holistic designs differentiated by design factors that were rooted in unique combinations of package design elements. To capture the type of holistic design, we labeled these groups as massive, contrasting, natural, delicate, and nondescript. An analysis of variance linked the prototypical holistic package designs to generic consumer responses.

Overall, the results provide the desired empirical support for holistic package designs rooted in unique combinations of design factors. They further indicate that package design conveys a strategically valued set of brand impressions, analogous to how consumers infer impressions from endorsers, advertising, and pricing. Associations between types of designs and the resultant generic brand impressions are distinctive, highlighting the holistic designs that firms, managers, and designers should consider for achieving specific responses. However, because it could be argued that our core findings apply only to the example product category, we extend the research to generate additional evidence for the generalizability of our findings.

An Extension Beyond the Wine Product Category

To provide evidence that our research can be applied outside the wine category, we replicated core parts of the original approach with fragrances. We selected this product category because (1) it meets the original requirements for this type of research (i.e., package design elements common to many other categories, availability of unfamiliar brands, package design creating impressions and driving consumer choice, and large variance in designs and brand personalities), and (2) compared with wine, there are obvious differences in the product category. Thus, a meaningful extension of the research in the fragrances category will support the existence of key types of holistic package design and their systematic relationships to generalizable brand impressions.

Stimuli and Evaluations

In contrast to the original procedure, we did not identify individual package design elements and assemble a pool of representative designs. Instead, we obtained a comprehensive list of fragrances (eaux de toilettes) from a leading national retailer. From this list of 446 designs, we obtained a sample by randomly selecting 120 packages. Stimuli consisted of standardized high-resolution digital images provided by the retailer.

We integrated these images into a PowerPoint presentation for uniform display to ten professional designers who had not participated in any of the previous studies. The designers were paid to categorize each stimulus using the five holistic package design types. They were briefed to ensure that they understood the factors and elements underlying each holistic design (see Table 2), and they were asked to consider all options carefully before making a decision. If they were unable to assign a design to any of the five categories (i.e., massive, contrasting, natural, delicate, and nondescript), they were instructed to assign it to a sixth "too hard/no good fit" category. On average, this task took the designers just over one hour for the initial rating. In addition, each professional reevaluated the resultant categorizations one more time three days later to ensure the reliability of the ratings.

Overall, the designers found the holistic design types to be useful and appropriate for categorizing fragrances. With 74 stimuli (61.7%), all ten designers unanimously agreed on the appropriate category; contrasting designs represented the largest group (30.0%) and natural designs represented the smallest group (10.0%) in the sample. In 17 cases in which at least one designer disagreed, the design was assigned to the category agreed on by the majority of judges. Because the designers could easily assign all stimuli to one or the other holistic type, we retained all stimuli for further analyses.

To measure design-evoked impressions, we conducted a survey with a convenience sample of consumers who had not participated in the previous studies. Panel members were invited and paid to view the stimuli the designers evaluated. As previously, participants controlled the speed of viewing and completed the 15-item brand personality scale and a measure of brand familiarity. Overall, we obtained 17,460 usable impressions ratings from 108 respondents.

Analyses and Results

In our analyses of the fragrances data, we replicated the procedures applied to the wine data. The unit of analyses was the individual package designs (N = 120), and the variables included 15 brand personality items (obtained from consumers) and the type of holistic design (obtained from professionals). In a few cases, average scores for the responses were based on only nine individual ratings because of data discarded as a result of consumer high familiarity with the brand. As previously, we analyzed brand personality data through CFA, producing five factors in line with Aaker's (1997) scale (CFI = .94). On the basis of these results, we averaged item scores to produce five universal brand personality dimensions. To assess the relationships between holistic package designs and the generalizable responses captured by the brand personality dimensions, we employed an analysis of variance. Table 6 shows significant differences in all response dimensions, as generated by the holistic design types; it also shows some examples. For fragrances, massive packages are associated with below-average impressions of excitement and sophistication and above-average ruggedness. Contrasting packages generate impressions of excitement along with high ruggedness and low competence. Natural designs appear sincere and sophisticated. Delicate designs are associated with competence, sophistication, and low ruggedness. Nondescript designs are associated with low sincerity, low excitement, and low ruggedness.

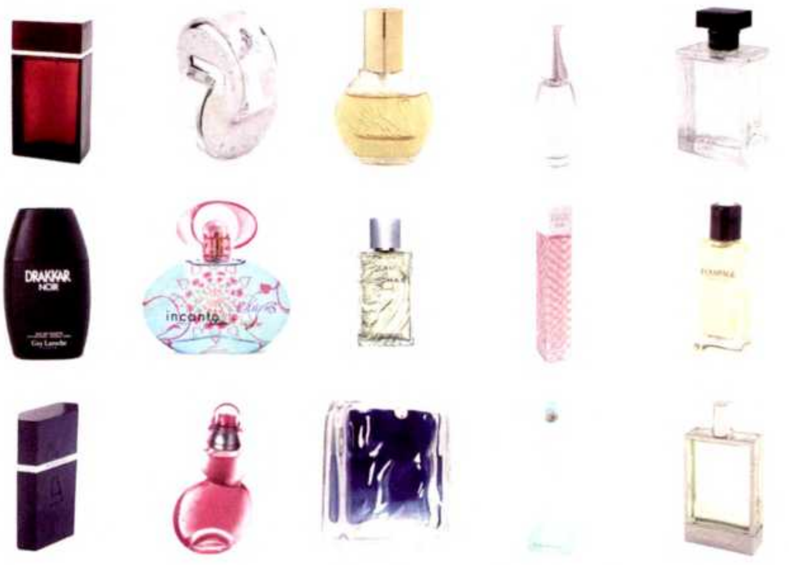
General Discussion

Managerial Implications: Guidelines for Developing and Modifying Package Designs

Note that the five types of holistic package designs may not be exhaustive (all inclusive), and other sources may exist for stimulating more specific brand impressions. However,

TABLE 6
Prototypical Holistic Package Designs and Impressions for Fragrances

	Holistic Design						
Response Dimension	**Massive (N = 27)**	**Contrasting (N = 36)**	**Natural (N = 12)**	**Delicate (N = 19)**	**Nondescript (N = 26)**	**Sample Mean**	***p***
Sincerity	2.17^{a}	2.11^{b}	**2.98**^{abc}	2.10	**2.03**^{c}	2.20	.001
Excitement	**2.53**^{abc}	**3.48**^{adef}	2.82^{bdg}	3.02^{ceh}	**2.40**^{fgh}	2.89	.001
Competence	2.88^{ab}	**2.62**^{acd}	2.75^{e}	**3.18**^{bcef}	2.87^{df}	2.84	.001
Sophistication	**2.51**^{abc}	2.95^{ad}	**3.07**^{bef}	**3.29**^{cdeg}	2.70^{fg}	2.86	.001
Ruggedness	**3.09**^{abc}	**2.91**^{de}	2.72^{afg}	**2.14**^{bdf}	**2.41**^{ceg}	2.70	.001



Notes: Values in bold distinguish the cluster in terms of emphasizing that dimension score to be significantly ($p < .05$) greater or smaller than the sample mean. Within each dimension, identical superscripts identify pairs of designs that score significantly ($p < .05$) different on this factor on the basis of a Scheffé post hoc test.

analogous to how the brand personality concept assists firms in brand positioning and differentiation, our holistic design types assist companies in more confidently using package design for conveying brand impressions.

First, managers can better communicate their package design needs using the taxonomy of the five types of holistic designs. Given the myriad design elements relevant in different contexts and settings and considering that few managers have design experience, it appears particularly useful that packages can be described by five holistic design types that plausibly exist for a wide range of products. Table 7 offers potential examples for massive, contrasting, natural, delicate, and nondescript package designs for varying product categories, including cereal, detergents, soft drinks, soups, and tea. Extending our research from packages to products, Table 7 also offers product design examples from other categories. We show prototypical designs for eyewear, casual shoes, MP3 players, and watches. In any case, the ability to describe design using a taxonomy of understandable holistic types is important because it enables managers to communicate better with designers. It provides the shared vocabulary, associations, and conventions that are so eloquently called for in previous research (McCracken 1986; Mick 1986).

Second, because our research shows that holistic package designs vary greatly in the generic response dimensions they create, companies can more accurately create brand personality impressions using holistic designs. Ultimately, design-evoked impressions should translate into gains in brand strength and equity. Initially, managers need to determine which impressions are desirable for their brand. Appropriate designs can then be created to achieve these desired responses. Our findings provide guidance on this issue by establishing systematic relationships between types of holistic package designs and generic response dimensions. Across most comparisons in Tables 5 and 6, the pattern of elicited responses (i.e., whether a design type generates below-average, above-average, or just average impressions) varied for fragrances and wine, thus indicating product category effects analogous to category-specific differences in types of brand personalities. A rugged wine brand is clearly related to impressions other than those for a rugged brand of fragrances. However, several design-response relationships appear to be robust across product categories. For example, massive designs are associated with low sophistication and average sincerity; contrasting designs generate high excitement, high ruggedness, low competence, and average sincerity; and natural designs convey sincerity and sophistication but only average ruggedness. Accordingly, managers interested in sincerity may want to consider natural designs because these tend to rate high on sincerity, though such designs do not guarantee perceived sincerity (e.g., if done poorly, if they interact negatively with some product categories), and there may be other, more specific means of stimulating perceived sincerity, such as a photo of the company's founder or family members. Given the occasional design–message variance, in some circumstances, the brand personality dimension communicated by a given holistic design might vary across product categories, even though the message is likely to be constant in most cases. Given consumer systematic responses to holistic designs, marketers should adopt a proactive approach and consider how they want consumers to perceive a brand. Rather than leaving impressions to chance, research with target consumers using preproduction package prototypes or illustrations can then determine whether a design will actually evoke the desired impressions.

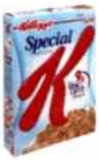
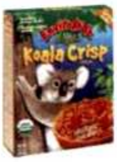
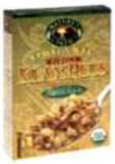
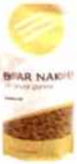
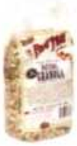
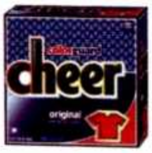
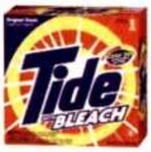
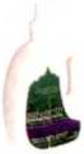
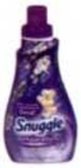
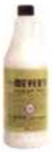
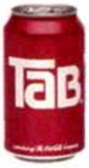
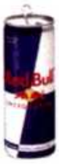
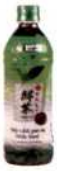
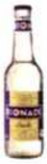
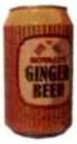
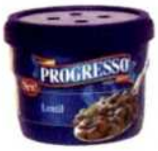
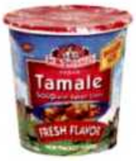
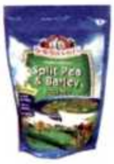
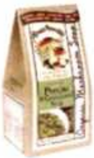
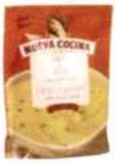
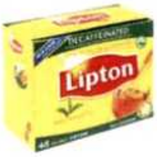
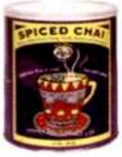
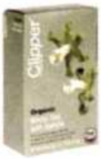
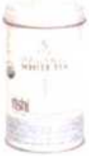
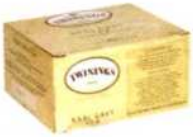
Third, despite the previously discussed trade-offs in impressions evoked by a holistic design type, designers have flexibility in creating appropriate packages. Our findings that holistic designs are characterized by generalizable design factors that, in turn, are rooted in specific design elements enable designers to achieve managerial objectives through various means, allowing them significant latitude for creativity. For example, Tables 3 and 4 show that a design could be massive because it has large, vertically oriented labels; large logos and images; and expansive surface areas. A wine package does not need to have all these characteristics to be perceived as massive, meaning that designers and managers have considerable leeway in how they accomplish such a holistic package design.

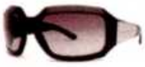
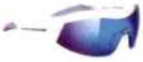
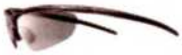
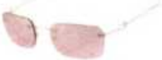
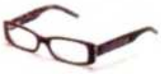
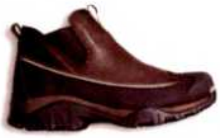
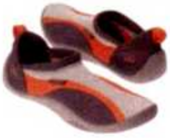
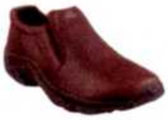
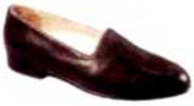
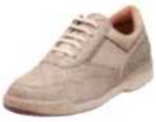
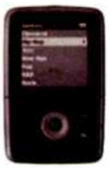
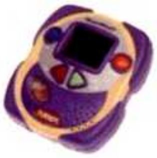
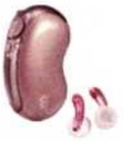
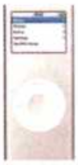
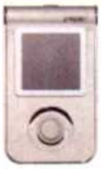
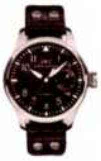
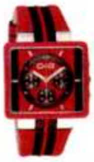
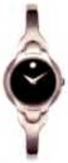
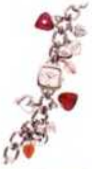
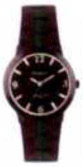
Finally, our findings indicate that holistic designs are useful in identifying visual competitors; packages within a holistic design type appear similar to consumers. Thus, managerial insight into which design group a firm's package and competing packages fall will improve brand positioning and package design selection. Related to the question of differentiating versus copying brands, prior research shows that clusters of similar alternatives commonly increase choice of alternatives of this type, according to what is called (positive) cluster effects or (negative) lone-alternative effects (Glazer, Kahn, and Moore 1991; Sivakumar 1995). Under some circumstances, however, cluster effects are negative (Brenner, Rottenstreich, and Sood 1999). Sood, Rottenstreich, and Brenner (2004) reconcile the effects by showing that cluster effects are positive when a person simply chooses an alternative from the available set but are negative when a person first chooses his or her preferred option within the clustered set and then chooses between that alternative and the otherwise lone alternative. Accordingly, clusters of similar package designs (visual competitors) might produce a general negative cluster effect and/or a positive lone-alternative effect, such that products perceived as being the same on substantive dimensions but offer unique designs are more likely to be chosen because of their unique packages (i.e., uniqueness on brand product features might be more of a general negative than uniqueness on design dimensions).

Theoretical Implications

By identifying prototypical holistic package designs and their underlying design factors and by establishing relationships between design types and generic response dimensions, this research makes several contributions to the literature on design and brand management. First, in line with Gestalt psychology, design processing theory, and empirical research on package design elements, we confirm that design elements are perceived as and organized into more complex design dimensions. The existence of such general-

TABLE 7
Examples from Other Product Categories

	Holistic Design				
Product Category	**Massive**	**Contrasting**	**Natural**	**Delicate**	**Nondescript**
Cereal					
Detergents					
Soft drinks					
Soups					
Tea					

	Holistic Design				
Extension to Product Designs	**Massive**	**Contrasting**	**Natural**	**Delicate**	**Nondescript**
Eyewear					
Casual shoes					
MP3 players					
Watches					

izable design factors that capture the essence of underlying design characteristics has been postulated in design processing theories (Geistfeld, Sproles, and Badenhop 1977; Veryzer 1999) and has been empirically verified in studies on selected package design elements, such as logos and typeface (Henderson and Cote 1998; Henderson et al. 2003; Henderson, Giese, and Cote 2004). However, our study is the first to assemble a universe of package design elements systematically and to trace how prototypical holistic package designs are rooted in generalizable design factors consisting of these elements.

Second, this research extends prior theorizing and empirical work by identifying not only factors of design but also holistic package designs as key types of visual stimuli representing generic classes or categories. The concept of holistic package design is important from the perspectives of both Gestalt psychology (Wertheimer 1925) and perceptual fluency (Janiszewski and Meyvis 2001). In addition, we add a new perspective to the discussion on part–whole distinctions. A person may comprehend the form of a package as a complete entity or as a collection of shapes, colors, typography, images, and add-ons. Durgee (1988) suggests that reactions to forms are based on atomistic perceptions. Accordingly, consumers might attend to individual design elements and the fit among them. Complex designs and those with conflicting elements tend to elicit the most elaborate cognitive processing (Schoormans and Robben 1997). For example, seeing the contrasting design of a Scion xB might inspire consumers to try to understand which characteristics might be present in this car, which is sharp edged and different from most other cars. Although our research does not actually demonstrate the independence of part–whole perception, it suggests that the role of Gestalt processing is important in how consumers perceive and respond to package designs. The package may first be perceived as a whole, and if the initial evaluation of the holistic design warrants further processing, individual design elements may become salient. Thus, a consumer encountering an Apple iPod might first consider the object in its entirety. Consumers who find the design sufficiently engaging might further attempt to analyze the appearance of the MP3 player. They might then process specific design elements, such as silhouette, scale, color scheme, and proportions, individually when contemplating the design.

Finally, the evidence we present strongly argues for the importance of examining package design as a means for generating brand impressions. To date, anecdotal evidence suggests that designers choose visual elements to generate brand associations, and the literature has documented fragmented evidence for several specific impressions generated by a particular package (Batra and Homer 2004; Forty 1992; Underwood 2003). Our research advances theory by demonstrating that prototypical holistic designs are systematically related to generalizable response dimensions. We are not aware of any study that has attempted to examine this issue of relationships between holistic package design and generic responses, probably because of the complexity involved. Although our results support the links identified in existing research, they also highlight the limitations of underlying theories. In particular, none of the current design theories are fully supported by our results because they do not extend beyond the identification of specific design elements or factors as contributors to brand impressions. No aesthetics or design theory considers more than one of the dimensions of design, and none address the importance of holistic design types. These theories also do not consider generalizable responses. As our results show, the influence of specific design elements and factors varies by holistic type. This implies a more complex relationship between design and responses than previously reported. Thus, no existing theory appears to be adequate for explaining relationships between design dimensions and consumer responses. Consistent with prior work (Henderson, Giese, and Cote 2004), this recognition is especially important because most marketing research is theory driven, but empirical results suggest that there is no good theory when it comes to aesthetics.

Study Limitations and Further Research

Although our research offers valuable implications for researchers and practitioners, it also has limitations that merit attention. Most important, we examine two types of products: wines and fragrances. Studying design is intrinsically difficult because important design elements and dimensions depend to some extent on the product and context being studied. It cannot be excluded that effect sizes change for products with vastly different package designs (e.g., packages of toothpaste or mouthwash may be more important in the purchase decision than a shoe box) or that differences occur as a result of different relative saliencies of brand impressions (e.g., the role of functional, symbolic, and experiential benefits in brand choice). However, prior research has identified generalizable design elements, factors, and relationships to responses that extend beyond a specific context (Henderson and Cote 1998; Henderson et al. 2003; Henderson, Giese, and Cote 2004). We addressed this context limitation throughout the article, specifically by showing that generalizable holistic package designs are related to generalizable response dimensions. The products chosen for our research might reduce effect sizes and impose limitations on how far the guidelines can be extended. However, these limitations should not be more significant than they are in any other study in which consumer responses are examined with a few examples and conclusions are extended to many types of products (e.g., when a single product is used for examining advertising elements).

Along similar lines, external validity concerns require more attention. Although our study improves on prior research through use of a controlled environment and consumer samples, a better understanding of how the focal variables interact with other variables is necessary. Findings from our single-country setting and a specific set of consumers may vary as a result of cultural context (Henderson et al. 2003) and centrality of visual stimulation versus other senses (Bloch, Brunel, and Arnold 2003). First, our research employed a sample of U.S. consumers. Although the sample was diverse in terms of ethnicity, age, and socioeconomic status, employing a sample with a different cultural background may result in different outcomes. Prior research has established that cultural dimensions influence the way people perceive and respond to different forms of design, particularly the brand impressions formed on the basis of the package (Aaker, Benet-Martinez, and Garolera 2001). Thus, a person's cultural background may influence the responses generated by a holistic package design, as well as other responses relevant to marketers.

Second, further research is needed to improve the understanding of a design's influence on consumer responses. Following approaches established in experimental aesthetics (Seifert 1992), we conducted analyses at the stimulus level rather than at the consumer level. Future studies could apply and extend the findings to include personality traits, particularly Bloch, Brunel, and Arnold's (2003) centrality of visual product aesthetics (CVPA). Thus, consumer segments could be identified according to their level of response to package design. Additional research is needed to determine the impact of CVPA on the extent of impression transfer from package to brand and its impact on the various responses and behaviors studied in marketing. Follow-up studies could replicate the analyses for groups of people who score either low or high on the CVPA scale. It is possible that individual differences occur in responses to holistic package designs and, consequently, differences in impressions evoked.

Third, several of the design elements and dimensions we included in this study not only influence brand impressions but simultaneously affect processing fluency (Reber, Schwarz, and Winkielman 2004), thus influencing brand recognition and recall (Janiszweski and Meyvis 2001). Future studies could determine the relationships between design-related processing fluency and brand impressions to develop guidelines for creating package designs that not only evoke generalizable brand impressions but also are high in visual fluency, thus aiding brand recognition and recall during consumer purchase processes.

Finally, the notion that design adds value to products implies that some consumers may be willing to pay a price premium for appealing designs (Bloch 1995; Borja de Mozota 2003). Brand impressions such as value for money, inexpensive, prestigious, and high quality appear to be closely related to perceptions of monetary value. Given previous findings that consumers develop price expectations for a product and then use this price to evaluate the product (Jun, MacInnis, and Park 2005), further research could investigate how package design influences consumer price perceptions. Specifically, information on which design dimensions contribute to price discrepancies (upward or downward differences between the prices consumers expect to pay versus the actual price) would help marketers better match design-evoked price expectations to actual prices, thus avoiding unfavorable expectation–disconfirmation effects.

In summary, this research speaks widely to the diverse community involved in package design, execution, and consumption. Although our findings might raise a new set of questions, we believe that meaningful answers have been provided, benefiting both future design activities and research.

REFERENCES

Aaker, David A. (1991), *Managing Brand Equity*. New York: The Free Press.

Aaker, Jennifer L. (1997), "Dimensions of Brand Personality," *Journal of Marketing Research*, 34 (August), 347–56.

——— (1999), "Brand Personality: A Path to Differentiation," in *Brands Face the Future*, Rory Morgan, ed. New York: Research International, 13–21.

———, Veronica Benet-Martinez, and Jordi Garolera (2001), "Consumption Symbols as Carriers of Culture: A Study of Japanese and Spanish Brand Personality Constructs," *Journal of Personality and Social Psychology*, 81 (3), 492–508.

———, Susan Fournier, and Brasel S. Adam (2004), "When Good Brands Do Bad," *Journal of Consumer Research*, 31 (1), 1–16.

Batra, Rajeev and Pamela Miles Homer (2004), "The Situational Impact on Brand Image Beliefs," *Journal of Consumer Psychology*, 14 (3), 318–30.

———, D.R. Lehmann, and D. Singh (1993), "The Brand Personality Component of Brand Goodwill: Some Antecedents and Consequences," in *Brand Equity and Advertising*, D.A. Aaker and A. Biel, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 83–96.

Belk, Russell W. (1988), "Possessions and the Extended Self," *Journal of Consumer Research*, 15 (2), 139–60

Berkowitz, Marvin (1987), "The Influence of Shape on Product Preferences," in *Advances in Consumer Research*, Vol. 14, Melanie Wallendorf and Paul F. Anderson, eds. Provo, UT. Association for Consumer Research, 573.

Berlyne, D.E. (1971), *Aesthetics and Psychobiology*. New York: Appleton-Century-Crofts.

Bloch, Peter H. (1995), "Seeking the Ideal Form: Product Design and Consumer Response," *Journal of Marketing*, 59 (July), 16–29.

———, Frederic F. Brunel, and Todd J. Arnold (2003), "Individual Differences in the Centrality of Visual Product Aesthetics: Concept and Measurement," *Journal of Consumer Research*, 29 (March), 551–65.

Borja de Mozota, Brigitte (2003), *Design Management: Using Design to Build Brand Value and Corporate Innovation*. New York: Allworth Press.

Brenner, Lyle, Yuval Rottenstreich, and Sanjay Sood (1999), "Comparison, Grouping, and Preference," *Psychological Science*, 10 (May), 225–29.

Caldewey, Jeffrey and Chuck House (2003), *Icon: Art of the Wine Label*. San Francisco: The Wine Appreciation Guild.

Caputo, Tina (2005), "The Repackaging of Fetzer," *Wines & Vines*, 86 (5), 20–25.

Carpenter, Patricia A. and Akira Miyake (1995), "Language Comprehension: Sentence and Discourse Processing," *Annual Review of Psychology*, 46, 91–120.

Durgee, Jeffrey F. (1988), "Product Drama," *Journal of Advertising*, 17 (February–March), 42–49.

Elliott, Richard L. and Kritsadarat Wattanasuwan (1998), "Brands as Symbolic Resources for the Construction of Identity," *International Journal of Advertising*, 17 (2), 131–44.

Folkes, Valerie and Shashi Matta (2004), "The Effect of Package Shape on Consumers' Judgments of Product Volume: Attention as a Mental Contaminant," *Journal of Consumer Research*, 31 (2), 390–401.

Forty, Adrian (1992), *Objects of Desire*. London: Thames and Hudson.

Fournier, Susan (1998), "Consumers and Their Brands: Developing Relationship Theory in Consumer Research," *Journal of Consumer Research*, 24 (4), 343–73.

Garber, Lawrence L., Raymond R. Burke, and J. Morgan Jones (2000), "The Role of Package Color in Consumer Purchase Consideration and Choice," Marketing Science Institute Report No. 00-104.

———, Eva M. Hyatt, and Richard G. Starr (2000), "The Effects of Food Color on Perceived Flavor," *Journal of Marketing Theory and Practice*, 8 (4), 59–72.

Geistfeld, Loren V., George B. Sproles, and Suzanne B. Badenhop (1977), "The Concept and Measurement of a Hierarchy of Product Characteristics," in *Advances in Consumer Research*, Vol. 4, William D. Perreault Jr., ed. Atlanta: Association for Consumer Research, 302–307.

Glazer, Rashi, Barbara E. Kahn, and William L. Moore (1991), "The Influence of External Constraints on Brand Choice: The Lone-Alternative Effect," *Journal of Consumer Research*, 18 (June), 119–27.

Grammer, Karl and Randy Thornhill (1994), "Human (*Homo sapiens*) Facial Attractiveness and Sexual Selection: The Role of Symmetry and Averageness," *Journal of Comparative Psychology*, 108 (3), 233–42.

Henderson, Pamela W. and Joseph A. Cote (1998), "Guidelines for Selecting or Modifying Logos," *Journal of Marketing*, 62 (April), 14–30.

———, ———, Siew Meng Leong, and Bernd Schmitt (2003), "Building Strong Brands in Asia: Selecting the Visual Components of Image to Maximize Brand Strength," *International Journal of Research in Marketing*, 20 (4), 297–313.

———, Joan L. Giese, and Joseph A. Cote (2004), "Impression Management Using Typeface Design," *Journal of Marketing*, 68 (October), 60–72.

Hertenstein, Julie H., Marjorie B. Platt, and Robert W. Veryzer (2005), "The Impact of Industrial Design Effectiveness on Corporate Financial Performance," *Journal of Product Innovation Management*, 2 (1), 3–21

Hirschman, Elisabeth (1986), "The Effect of Verbal and Pictorial Advertising Stimuli on Aesthetic, Utilitarian and Familiarity Perceptions," *Journal of Advertising*, 15 (2), 27–34.

Holbrook, Morris B. (1986), "Aims, Concepts, and Methods for the Representation of Individual Differences in Esthetic Responses to Design Features," *Journal of Consumer Research*, 13 (3), 337–47.

Janiszewski, Chris and Tom Meyvis (2001), "Effects of Brand Logo Complexity, Repetition, and Spacing on Processing Fluency and Judgment," *Journal of Consumer Research*, 28 (1), 18–32.

Johar, Venkataramani Gita, Jaideep Sengupta, and Jennifer L. Aaker (2005), "Two Roads to Updating Brand Personality Impressions: Trait Versus Evaluative Inferencing," *Journal of Marketing Research*, 72 (November), 458–69.

Jun, Sung Youl, Deborah J. MacInnis, and C. Whan Park (2005), "Formation of Price Expectation in Brand Extensions and Impact on Brand Extension Evaluation," in *Advances in Consumer Research*, Vol. 32, Geeta Menon and Akshay Rao, eds. Valdosta, GA: Association for Consumer Research, 137–42.

Keller, Kevin L. (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1–22.

——— (2003), "Brand Synthesis: The Multidimensionality of Brand Knowledge," *Journal of Consumer Research*, 29 (March), 595–600.

Koffka, Kurt (1922), "Perception: An Introduction to the Gestalt-Theories," *Psychological Bulletin*, 19 (October), 531–85.

Kunst-Wilson, W.R and R.B. Zajonc (1980), "Affective Discrimination of Stimuli that Cannot Be Recognized," *Science*, 207 (4430), 557–58.

Lawson, Bryan (1983), *How Designers Think*. Westfield, NJ: Eastview Editions.

Lewicki, Pawel (1986), *Nonconscious Information Processing*. New York: Academic Press.

Loken, Barbara and James Ward (1990), "Alternative Approaches to Understanding the Determinants of Typicality," *Journal of Consumer Research*, 17 (September), 111–26.

Mackay, Jordan (2005), "Labels Should Convey Both Style and Substance," *Wines & Vines*, 86 (5), 96.

McCracken, Grant (1986), "Culture and Consumption: A Theoretical Account of the Structure and Movement of Cultural Meaning of Consumer Goods," *Journal of Consumer Research*, 13 (1), 71–84.

Mick, David Glen (1986), "Consumer Research and Semiotics: Exploring the Morphology of Signs, Symbols, and Significance," *Journal of Consumer Research*, 13 (2), 196–212.

Orth, Ulrich R. (2005), "Consumer Personality and Other Factors in Situational Brand Choice Variation," *Journal of Brand Management*, 13 (2), 115–33.

———, Marianne McGarry Wolf, and Tim Dodds (2005), "Dimensions of Wine Region Equity and Their Impact on Consumer Preferences," *Journal of Product and Brand Management*, 14 (2), 88–97.

Pepper, S.C. (1949), *Principles of Art Appreciation*. New York: Harcourt, Brace and World.

Quester, Pascale G. and Justin Smart (1998), "The Influence of Consumption Situation and Product Involvement Over Consumers' Use of Product Attributes," *Journal of Consumer Marketing*, 15 (3), 220–38.

Raghubir, Priya and Eric A. Greenleaf (2006), "Ratios in Proportion: What Should the Shape of the Package Be?" *Journal of Marketing*, 70 (April), 95–107.

——— and Aradhna Krishna (1993), "Vital Dimensions in Volume Perception: Can the Eye Fool the Stomach?" *Journal of Marketing Research*, 36 (August), 313–26.

Reber, Rolf, Norbert Schwarz, and Piotr Winkielman (2004), "Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience?" *Personality and Social Psychology Review*, 8 (4), 364–82.

Rettie, Ruth and Carol Bruwer (2000), "The Verbal and Visual Components of Package Design," *Journal of Product and Brand Management*, 9 (1), 56–70.

Schmitt, Bernd H. and Alex Simonson (1995), "Managing Corporate Image and Identity," *Long Range Planning*, 28 (5), 82–92.

Schoormans, Jan P.L. and Henry S.J. Robben (1997), "The Effect of New Package Design on Product Attention, Categorization and Evaluation," *Journal of Economic Psychology*, 18 (2–3), 271–87.

Seifert, Lauren Sue (1992), "Experimental Aesthetics: Implications for Aesthetic Education of Naïve Art Observers," *Journal of Psychology*, 126 (1), 73–78.

Sirgy, M. Joseph (1982), "Self-Concept in Consumer Behavior: A Critical Review," *Journal of Consumer Research*, 9 (3), 287–300.

Sivakumar, K. (1995), "Role of Price and Quality Tiers on the Cluster Effect in Brand Choice," *Marketing Letters*, 6 (October), 265–73.

Sood, Sanjay, Yuval Rottenstreich, and Lyle Brenner (2004), "On Decisions that Lead to Decisions: Direct and Derived Evaluations of Preference," *Journal of Consumer Research*, 31 (June), 17–25.

Sung, Yongjun and Spencer F. Tinkham (2005), "Brand Personality Structures in the United States and Korea: Common and Culture-Specific Factors," *Journal of Consumer Psychology*, 15 (4), 334–50.

Sweeney, Jillian C. and Carol Brandon (2006), "Brand Personality: Exploring the Potential to Move from Factor-Analytical to Circumplex Models," *Psychology & Marketing*, 23 (8), 639–63.

Underwood, Robert L. (2003), "The Communicative Power of Product Packaging: Creating Brand Identity via Lived and Mediated Experience," *Journal of Marketing Theory and Practice*, 9 (Winter), 62–76.

——— and Noreen M. Klein (2002), "Packaging as Brand Communication: Effects of Product Pictures on Consumer Responses to the Package and Brand," *Journal of Marketing Theory and Practice*, 10 (4), 58–68.

——— and Julie L. Ozanne (1998), "Is Your Package an Effective Communicator? A Normative Framework for Increasing the Communicative Competence of Packaging," *Journal of Marketing Communication*, 4 (4), 207–220.

Vernadakis, George (2000), "What the 'Bug' Can Teach Gates: Brand Personality Goes a Long Way," *Advertising Age*, 71 (27), 48.

Veryzer, Robert W. (1999), "A Nonconscious Processing Explanation of Consumer Response to Product Design," *Psychology and Marketing*, 16 (6), 497–522.

——— and J. Wesley Hutchinson (1998), "The Influence of Unity and Prototypicality on Aesthetic Responses to New Product Designs," *Journal of Consumer Research*, 24 (March), 374–94.

Wansink, Brian (1996), "Can Package Size Accelerate Usage Volume?" *Journal of Marketing*, 60 (July), 1–14.

Wertheimer, Max (1925), "Untersuchungen zur Lehre von der Gestalt: II (Investigations in Gestalt Theory: II)," *Psychologische Forschung*, 4 (2), 301–350.

Yang, Sha and Priya Raghubir (2005), "Can Bottles Speak Volumes? The Effect of Package Shape on How Much to Buy," *Journal of Retailing*, 81 (4), 269–81.

Markus Wübben & Florian v. Wangenheim

Instant Customer Base Analysis: Managerial Heuristics Often "Get It Right"

Recently, academics have shown interest and enthusiasm in the development and implementation of stochastic customer base analysis models, such as the Pareto/NBD model and the BG/NBD model. Using the information these models provide, customer managers should be able to (1) distinguish active customers from inactive customers, (2) generate transaction forecasts for individual customers and determine future best customers, and (3) predict the purchase volume of the entire customer base. However, there is also a growing frustration among academics insofar as these models have not found their way into wide managerial application. To present arguments in favor of or against the use of these models in practice, the authors compare the quality of these models when applied to managerial decision making with the simple heuristics that firms typically use. The authors find that the simple heuristics perform at least as well as the stochastic models with regard to all managerially relevant areas, except for predictions regarding future purchases at the overall customer base level. The authors conclude that in their current state, stochastic customer base analysis models should be implemented in managerial practice with much care. Furthermore, they identify areas for improvement to make these models managerially more useful.

Keywords: customer base analysis, heuristics, relationship marketing, direct marketing, customer management

Consider a marketing executive at a catalog retailer who faces the following challenges: First, she wants to distinguish customers in the customer base who are likely to continue buying from the firm (active customers) from those who are likely to defect or from those who have already defected (inactive customers). This information should help (1) identify profitable, inactive customers who should be reactivated; (2) remove inactive, unprofitable customers from the customer base; and (3) determine active customers who should be targeted with regular marketing activities, such as new catalogs or mailings. Second, she wants to generate transaction forecasts for individual customers to identify the company's future 10% best customers, or to compute customer lifetime value (CLV). Such information should help her target those groups with perks, differential mailing frequencies, and loyalty program offerings. Third, she wants to predict the purchase volume of the entire customer base to make provisions for capacity planning, to compute the firm's customer equity, and to know when customer acquisition efforts need to be strengthened.

For the executive, the central problem in successfully coping with these tasks is that the time at which a customer defects from the firm is unobservable. The customer may have been disenchanted with the purchased product or the provider and now buys at a different supplier, the customer may have moved to another city, or the customer may have even passed away. This phenomenon exists for most service providers that operate in noncontractual settings: For example, when a customer purchases from a catalog retailer, walks off an aircraft, checks out of a hotel, or leaves a retail outlet, the firm has no way of knowing whether and how often the customer will conduct business in the future (Reinartz and Kumar 2000).

In contrast, in a contractual setting, the buyer–seller relationship is governed by a contract, which often predetermines not only the length but also the usage pattern of the relationship (e.g., telephone and Internet "flat-rate" services, magazine subscriptions). In this context, hazard regression or logistic regression models (Bolton 1995; Li 1995) provide promising approaches in determining the probability that a customer will still be with the firm at a particular future time. In the noncontractual setting, the state-of-the-art approach in determining the activity and future purchase levels of a customer is the Pareto/NBD model (Schmittlein, Morrison, and Columbo 1987; Schmittlein and Peterson 1994). The Pareto/NBD model has recently been employed in several studies (Fader, Hardie, and Lee 2005a, b; Ho, Park, and Zhou 2006; Krafft 2002; Reinartz and Kumar 2000, 2003), and its implementation has been recommended on an even larger scale (Balasubra-

---

Markus Wübben is a doctoral student (e-mail: markus.wuebben@wi.tum.de), and Florian v. Wangenheim is Chair of Services and Technology Marketing (e-mail: florian.wangenheim@wi.tum.de), TUM Business School, Technische Universität München. The authors thank Katherine N. Lemon, Bruce Hardie, Pete Fader, Dan Goldstein, and the anonymous *JM* reviewers for their valuable comments during the preparation of this manuscript. This work was supported by the German Research Foundation (DFG).

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

manian et al. 1998; Jain and Singh 2002; Kamakura et al. 2005; Rust and Chung 2006). Recently, Fader, Hardie, and Lee (2005a) introduced the BG/NBD model, which is a variant of the Pareto/NBD model but is much easier to implement and estimate. Both models are attractive because they (1) make forecasts of individuals' future purchase levels and (2) operate on past transaction behavior. More precisely, they operate solely on the frequency and recency information of a customer's past purchase behavior. The Pareto/NBD model has an additional feature; for each customer, it yields the probability that he or she is still active.

In light of increased calls for closer cooperation between marketing academics and practitioners, it must be of concern for academics that these models have not found their way into managerial practice. Instead, a survey by Verhoef and colleagues (2002) shows that simple heuristics are still commonly applied.

Given the time and money costs associated with implementing complex stochastic models in managerial practice, the marketing executive will be convinced to make use of the academic methods only when their superiority is clearly demonstrated on the aggregate level and, even more important, on the individual customer level. However, practitioners are not the only ones who would benefit from such insights. For research, it is important to know the circumstances under which the predictions of these models can be trusted to produce accurate forecasts for future implementation of these models in, for example, CLV research (e.g., Reinartz and Kumar 2000, 2003).

Few studies have compared the performance of complex versus noncomplex models for customer purchase behavior and lifetime value prediction. Donkers, Verhoef, and De Jong (2007) find that using complex methods instead of simple models for CLV prediction in a contractual setting (insurance company) does not substantially improve predictive accuracy. In a semicontractual context, Borle, Singh, and Jain (2008) find that a simple RFM (recency, frequency, and monetary value) model performs as well as the Pareto/NBD model that includes monetary value (Schmittlein and Peterson 1994) in predicting CLV. They also propose a hierarchical Bayesian model that works better than both the Pareto/NBD and the RFM models in the semicontractual setting. However, none of the studies on the stochastic models (Fader, Hardie, and Lee 2005a; Schmittlein, Morrison, and Columbo 1987; Schmittlein and Peterson 1994) have validated their predictions on the individual customer level in a noncontractual setting using multiple data sets from different industries. The current research aims to fill this gap.

In what follows, we briefly cover heuristics in managerial practice and provide an introduction to the Pareto/NBD and BG/NBD models. We then describe the data sets from three different industries on which we performed our validation. Next, we present the results of the predictive performance of the models versus the simple management heuristics. Finally, we discuss our findings and offer recommendations for using customer base analysis models in academic research and managerial practice.

Heuristics in Managerial Practice

Heuristics are often mentioned in managerial literature on customer relationship management or database marketing (e.g., Hughes 2006; Novo 2004). For example, Blattberg, Getz, and Thomas (2001) suggest the recency–sales matrix, which is a slightly modified variant of the hiatus heuristic. The most prominent example of a heuristic in practical customer management is probably the RFM framework, which was introduced by Alden's catalog company in the 1920s to value customers to decide which customer should receive a catalog (Roel 1988).

The diffusion of simple heuristics can be explained by the opportunity costs that managers experience when investing time and effort in managerial decision making. Rieskamp and Hoffrage (2008) show that under opportunity costs, people tend to make decisions quickly. A framework for explaining contingent decision behavior—in this case, choosing between complex methods and simple heuristics—focuses on the accuracy and cognitive effort characterizing the available strategies (Payne, Bettman, and Johnson 1993). The basic hypothesis of the effort/accuracy framework is that the strategy used to make a decision represents a balancing of the goals of being as accurate as possible and conserving limited cognitive resources. Creyer, Bettman, and Payne (1990) find that when the goal is to maximize accuracy rather than minimize effort, people acquire more information, take more time, are more alternative based, and are ultimately more accurate. However, given opportunity costs, managers aim to minimize time and effort and therefore resort to well-established, simple heuristics. As Richard Abdoo, chair and chief executive officer of Wisconsin Energy Corporation, notes, "As we move to a deregulated marketplace, we don't have this slow process of hearings and review and two years to make a decision. We now have to make decisions in a timely manner. And that means that we process the best information that's available and infer from it and use our intuition to make a decision" (Hayashi 2001, p. 61).

There are reasons to believe that simple heuristics may work better than more complex strategies for various types of tasks, even though they often require less information and computation (Gigerenzer, Todd, and The ABC Research Group 1999). Support for Grigerenzer's work is manifold (Bröder 2000, 2003; Bröder and Schiffer 2003; Lee and Cummins 2004; Newell et al. 2004; Newell and Shanks 2003; Newell, Weston, and Shanks 2003; Rieskamp 2006; Rieskamp and Hoffrage 1999; Rieskamp and Otto 2006). A survey conducted by Jagdish Parikh (discussed in Buchanan and O'Connell 2006, p. 40) shows that executives "used their intuitive skills as much as they used their analytical abilities, but they credited 80% of their successes to instinct." Especially in direct marketing, experienced managers are likely to be accurate (Morwitz and Schmittlein 1998) because this environment is characterized by two properties that are essential for learning to occur: repetition and feedback (Camerer and Johnson 1991; Goldberg 1968). "Repetition" refers to the repeated occurrence of the task, and "feedback" means that the outcome of the manager's decision is easily observed and evaluated. The previously described decisions to distinguish between active and inactive customers and between high- and low-value customers both are highly repetitive and offer feedback.

A survey conducted in May 2002 by executive search firm Christian and Timbers reveals that 45% of corporate executives now rely more on instinct than on facts and figures in running their businesses (Bonabeau 2003). Nevertheless, even if facts and figures are used, managers still rely on intuition and long-standing methods. A survey by Verhoef and colleagues (2002) on 228 database marketing companies shows that cross-tabulation and RFM analysis are the most popular methods for response modeling. In the context of the current study, at least two of the three companies (airline and apparel retailer) whose customer bases are analyzed apply simple recency-of-last-purchase (hiatus) analysis to distinguish active from inactive customers. For example, the managers of the airline who have "expert" knowledge of their customer base informed us that the cutoff time was nine months. This finding is in line with an article in the *New York Times* (Wade 1988) that illustrates the use of the hiatus heuristic in frequent-flier programs. For future purchase–level determination, average past purchase behavior is often employed as a simple predictor for future behavior. The managers of focal firms also confirmed this. A series of short telephone interviews with eight people responsible for customer management within their firms revealed that the hiatus heuristics was applied in all companies to determine active and inactive customers. For determining future best customers, somewhat varying approaches were applied, but number of past purchases was always a central variable (e.g., in an RFM-type approach).

Stochastic Customer Base Analysis Models

Both the Pareto/NBD and the BG/NBD models were developed to model repeat-buying behavior in a setting in which customers buy at a steady (albeit stochastic) rate and eventually become inactive at some unobserved time. The information they operate on consists solely of customers' past purchase behavior. More precisely, for each customer, the models operate on three values $(X = x, t, T)$, where $X = x$ is the number of purchases made in time frame $(0, T]$, with the last purchase occurring at time t, where $0 < t \leq T$. In addition, the models must be calibrated on the customer base to which they are applied. This calibration process yields several model parameters that describe the purchase and dropout process of the analyzed customer base.

Pareto/NBD

The Pareto/NBD model builds on the assumption that purchases follow Ehrenberg's (1988) NBD model, whereas dropout events follow a Pareto distribution of the second kind. More precisely, the Pareto/NBD model assumptions are as follows:

1. Individual Customer
   - Poisson purchases: While active, each customer makes purchases according to a Poisson process with purchase rate λ.
   - Exponential lifetime: Each customer remains active for a lifetime, which has an exponentially distributed duration with dropout rate μ.
2. Heterogeneity Across Customers
   - Individuals' purchase rates distributed gamma: The purchasing rate λ for the different customers is distributed according to a gamma distribution across the population of customers.
   - Individuals' dropout rates distributed gamma: The customers' dropout rates μ are distributed according to a gamma distribution across the population of customers.
   - Rates λ and μ are independent: The purchasing rates λ and the dropout rates μ are distributed independently of each other.

Among other things, the Pareto/NBD model yields the following information:

- $P(\text{Active}|X = x, t, T)$ is the probability that a random customer with purchase pattern $(X = x, t, T)$ (whose individual purchase rate and dropout rate may be unknown) is active at some time T (see Schmittlein, Morrison, and Colombo 1987, Equations 11, 12, and 13).
- $E(X^*|X = x, t, T, T^*)$ is the expected number of transactions X^* of a random customer with purchase pattern $(X = x, t, T)$ (and unknown individual purchase rate and dropout rate) in time $(T, T + T^*]$ (see Schmittlein, Morrison, and Colombo 1987, Equation 22).

BG/NBD

There is only one assumption in the BG/NBD model that differs from the assumptions of the Pareto/NBD model. Whereas the Pareto timing model assumes that dropout of a customer can occur anytime, the BG/NBD model assumes that dropout occurs only directly after purchases. This slight change greatly reduces the complexity of the model because a beta-geometric (BG) model can be used to represent the dropout phenomena instead of the exponential gamma (Pareto) model. More precisely, the BG/NBD model assumptions are as follows:

1. Individual Customer
   - Poisson purchases: While active, each customer makes purchases according to a Poisson process with purchase rate λ.
   - Geometric lifetime: Each customer remains active for a lifetime, which is distributed over the number of transactions according to a (shifted) geometric distribution with dropout probability p.
2. Heterogeneity Across Customers
   - Individuals' purchase rates distributed gamma: The purchase rate λ for the different customers is distributed according to a gamma distribution across the population of customers.
   - Individuals' dropout probabilities distributed beta: The customers' dropout probabilities p for different customers is distributed according to a beta distribution across the population of customers.
   - Rates λ and p are independent: The purchase rates λ and dropout probabilities p are distributed independently of each other.

Among other things, the BG/NBD model yields the following information:

- $E(X^*|X = x, t, T, T^*)$ is the expected number of transactions X^* of a random customer with purchase pattern ($X = x, t, T$) (and unknown individual purchase rate and dropout rate) in time $(T, T + T^*]$ (see Fader, Hardie, and Lee 2005a, Equation 10).

The BG/NBD model also includes the expression $P(\text{Active}| X = x, t, T)$ to compute the probability that a customer is active at some time T, but the application of this expression is limited to customers whose individual purchase rate λ and dropout probability p is known. However, determining the dropout probability p for an individual is virtually impossible.

Data

We conducted our study on three different data sets from three different industries. The first data set comes from an apparel retailer and covers 46,793 customers and their purchases from January 2003 through August 2004. We based our analysis on a cohort of 2330 customers who began their buyer–seller relationship with the apparel retailer in the last week of January 2003. Thus, for this cohort, the available data cover the initial and repeat purchases for each customer over a period of 80 weeks. To calibrate the models, we used repeat-purchase data for the 2330 customers over the first 40 weeks of the 80-week period, leaving a 40-week holdout period to validate the models.

The second data set comes from a major global airline and covers 146,961 customers and their purchases from January 1999 through December 2002. The available data only provided aggregated quarterly transactions for each customer and did not include the exact purchase dates. Our analysis of this data set focused on a cohort of 2891 customers who conducted their initial purchase from the airline in the first quarter of 1999. For this cohort, we chose a calibration period of eight quarters (January 1999–December 2000), leaving eight quarters for the holdout period (January 2001–December 2002).

The third data set covers customers of the online CD retailer CDNOW. The data track 23,570 customers and their purchases from January 1997 through June 1998 (78 weeks), all of whom initiated their first purchase at CDNOW in the first quarter of 1997. Fader and Hardie (2001) already used this data in multiple studies. More precisely, we used the 2357 customer cohort available on Bruce Hardie's Web site (see Fader, Hardie, and Lee 2005a). The calibration and holdout periods are 39 weeks each. Detailed descriptive statistics of all three data sets appear in Table 1.

Analysis

Given the lack of empirical analysis for the superiority of the considered academic methods in determining active customers and forecasting future purchase levels, the following analyses try to shed light on this open question. First, we analyze how well the hiatus heuristic, which is used by the managers of the firms whose customer bases we analyze in this article, performs in comparison with the Pareto/NBD P(Active) facility. Second, we analyze how well the Pareto/NBD and BG/NBD models forecast future purchase behavior for both the individual customer and the customer base as a whole. Although aggregated sales forecasts are important statistics in terms of, for example, capacity planning or customer equity computation, we specifically focus on the forecast performance for the individual customer. This stems from the notion that there must be a decent individual customer purchase–level forecast for proper computation of metrics, such as CLV (Reinartz and Kumar 2000, 2003) or customer value segment classification (e.g., gold, silver, and bronze segments). Picking up on this idea, not only do we present mere performance measures for individuals' forecasts, but in a third analysis, we

TABLE 1
Descriptive Statistics

	Airline	Apparel	CDNOW
Sample size (n)	2891	2330	2357
Available time frame	16 quarters	80 weeks	78 weeks
Time split (estimation/holdout[a])	8/8	40/40	39/39
Available time units	Quarters	Weeks/months/quarters	Weeks
Zero repeaters in estimation periods	193	371	1411
Zero repeaters in holdout periods	1376	395	1673
Zero repeaters in estimation and holdout periods	163	184	1218
Number of purchases in estimation periods	31,479	10,855	2457
Number of purchases in holdout periods	23,033	11,351	1882
Average number of purchases per customer in estimation periods (SD)	10.88/customer (15.988)	4.658/customer (5.412)	1.04/customer (2.190)
Average number of purchases per customer in holdout periods (SD)	7.967/customer (16.810)	4.871/customer (5.598)	.798/customer (2.057)
Average T (SD)	4.393 quarters (3.006)	25.15 weeks (14.21)	6.845 weeks (10.731)

[a]Holdout period length was varied from 1 to max(holdout periods) in the analyses.

also show how well the models perform in identifying a company's future 10% (20%) best customers.

Parameter Estimation

Both the Pareto/NBD and the BG/NBD models need to be calibrated on the customer base to which they are applied. The Pareto/NBD model has four parameters (r, α, s, β), where (r, α) represent the shape and scale parameters of the gamma distribution that determines the distribution of the purchase rates across individuals of the customer base and (s, β) represent the scale and shape parameters of the gamma distribution that determines the distribution of the dropout rates across individuals. The BG/NBD model holds four model parameters (r, α, a, b) as well, where (r, α) (as in the NBD/Pareto model) determine the shape and scale of the purchase rate gamma distribution and (a, b) represent the shape parameters of a beta distribution that determines the distribution of the dropout probabilities across individuals of the customer base. For both models, we used a maximum likelihood approach under MATLAB to estimate the model parameters. Tables 2 and 3 report each cohort's parameters for the Pareto/NBD and BG/NBD models.

The Pareto/NBD and BG/NBD model parameters computed for the airline and CDNOW data sets are reasonable. According to the Pareto/NBD model, an average airline customer initiates 1.9877 transactions per quarter and remains active for 7.83 quarters.[1] An average CDNOW customer initiates .0523 transactions per week (one purchase every 19.12 weeks) and remains active for 19.26 weeks. According to the BG/NBD model, an average airline customer initiates 2.110 transactions per quarter and remains active for 4.34 quarters.[2] An average CDNOW customer initiates .0549 transactions per week (one purchase each 18.21 weeks) and purchases from the company for 73.95 weeks.

For the apparel data set, both the Pareto/NBD and the BG/NBD models compute notable results. Although the purchasing rate of an average customer (.1190 purchases per week [one purchase every 8.40 weeks] for the Pareto/NBD model and .1198 purchases per week or one purchase every 8.34 weeks for the BG/NBD model) is reasonable, the lifetime of an average customer is exceptionally long. According to the Pareto/NBD model, an average customer remains active for 909.09 weeks (~17.48 years). According to the BG/NBD model, an average customer remains active for 684.2 weeks (~13.15 years). In other words, the models predict an average apparel customer to be ultimately loyal. This effect will be reflected in very high P(Active) values of the apparel customers.

Determining Active and Inactive Customers

If the Pareto/NBD model is used for customer activity determination, each customer's P(Active) value is computed on the basis of the customer's purchase pattern in the observation period (i.e., estimation period). However, for the continuous P(Active) values to be useful in managerial application, a cutoff threshold $c_{P(Active)}$ (decision boundary) must be determined. Customers whose P(Active) value is greater than or equal to $c_{P(Active)}$ are classified as active, and customers whose P(Active) value is less than $c_{P(Active)}$ are classified as inactive. More precisely, for the Pareto/NBD model, given a cutoff threshold $c_{P(Active)}$ and a customer

---

[1] Within the Pareto/NBD model, r/α represents the number of purchases of an average customer in one time unit, and s/β represents the dropout rate of an average customer per time unit. The lifetime of an average customer is exponentially distributed with parameter s/β and has an expected value of 1/(s/β). Therefore, according to the estimated parameters, an average CDNOW customer remains active for 1/(.0519) = 19.26 weeks.

[2] Within the BG/NBD model, r/α represents the number of purchases of an average customer in one time unit, and an average customer remains active until time τ, which is exponentially distributed with parameter pλ and has an expected value of 1/(pλ), given that λ = r/α and p = a/(a + b). Therefore, according to the estimated parameters, an average CDNOW customer remains active for 1/(.0549 × .2463) = 73.95 weeks.

TABLE 2
Results of the Pareto/NBD Maximum Likelihood Estimation

	r	α	r/α	s	β	s/β	Log-Likelihood
Apparel	1.0954	9.2029	.1190	1.0885	973.7829	.0011	–31338.7
Airline	1.4304	.7196	1.9877	2.5086	19.6408	.1277	–2150.2
CDNOW	.5533	10.5776	.0523	.6061	11.6650	.0519	–9595.0

TABLE 3
Results of the BG/NBD Maximum Likelihood Estimation

	r	α	r/α	a	b	a/(a + b)	Log-Likelihood
Apparel	1.0592	8.8371	.1198	.0324	2.6243	.0122	–31336.6
Airline	1.15186	.545781	2.11048	.456637	3.73439	.108956	–2238.34
CDNOW	.2426	4.4135	.0549	.7931	2.4260	.2463	–9582.43

with purchase pattern (X = x, t, T), the customer is classified according to the following:

- $P(\text{Active})_T \geq c_{P(\text{Active})} \Rightarrow$ Customer is classified as active, and
- $P(\text{Active})_T < c_{P(\text{Active})} \Rightarrow$ Customer is classified as inactive.

To validate the classifications, we use the holdout period according to the following scheme: If a customer has made at least one purchase in the holdout period, he or she is considered "active"; if the customer has not purchased in the holdout period, he or she is considered "inactive." This scheme induces four possible classification outcomes based on whether a customer has or has not been correctly classified as active or inactive.

Likewise, for the hiatus heuristic, there needs to be a cutoff threshold c_{hiatus} below which customers are classified as active and above which customers are classified as inactive. In other words, if a customer has not purchased for more than a time span of length c_{hiatus}, he or she is considered inactive; otherwise, he or she is considered active. More precisely, let (X = x, t, T) be a customer's purchase pattern and c_{hiatus} be a cutoff threshold. Then,

- $T - t < c_{\text{hiatus}} \Rightarrow$ Customer is classified as active, and
- $T - t \geq c_{\text{hiatus}} \Rightarrow$ Customer is classified as inactive.

We obtain the same four possible classification outcomes as in the P(Active) case.

For the Pareto/NBD model, a "natural" choice for the cutoff threshold $c_{P(\text{Active})}$ could be .5, which is in line with the work of Reinartz and Kumar (2000) on the Pareto/NBD model and the classification literature (Sharma 1996). Helsen and Schmittlein (1993) also use .5 in the prediction of purchase events in survival analysis.

For the hiatus heuristic, airline and apparel firm mangers informed us that customers were considered inactive if they had not purchased from the firm for more than nine months. For the CDNOW data set, we did not have access to this information. Furthermore, we do not know whether CDNOW uses the hiatus heuristic at all. Given that online firms operate in fast-moving markets, we decided to use a hiatus length of six months, which should match the circumstances of an online retailer.

However, neither the managers' chosen hiatus nor a P(Active) threshold of .5 may necessarily be optimal thresholds in terms of overall correctly classified customers. We observe how far the managers' chosen hiatus and the P(Active) threshold of .5 deviate from their "optimal" values and the effect of this difference in terms of the classification performance. First, we show how well the hiatus heuristic with the threshold determined by managers' expert knowledge distinguishes the active from the inactive customers in comparison with a P(Active) analysis with a natural cutoff threshold of .5. Second, we show how sensitive the classification performance is to the choice of the thresholds.

Table 4 shows the results of the first analysis. The hiatus heuristic performs better (in terms of overall correctly classified customers) than the P(Active) facility in two of the three cases; the P(Active) facility performs only slightly better on the CDNOW data set. Even more notable, the P(Active) facility fails to classify any of the inactive customers in the apparel data set correctly, whereas the hiatus heuristic classifies 47.84% of the inactive customers in the cohort correctly. This suggests that the optimal cutoff threshold $c_{P(\text{Active})}$ for this cohort may deviate considerably from .5. What about the other P(Active) cutoff thresholds $c_{P(\text{Active})}$ and hiatus heuristic cutoff thresholds c_{hiatus}? Are these optimal? If not, what are the optimal values? This is the subject of our next analysis.

As we mentioned previously, we consider a cutoff threshold of $c_{P(\text{Active})}$ or c_{hiatus} optimal if it maximizes the percentage of overall correctly classified active and inactive customers of a cohort using our classification procedure.[3] Our algorithm for finding the optimal cutoff thresholds simply iterates over the domain of valid cutoff thresholds. More precisely, for the Pareto/NBD model, the algorithm is as follows:

TABLE 4
P(Active) Versus Hiatus Heuristic

	Airline		Apparel		CDNOW	
	Three Quarters Hiatus	$P(\text{Active})_{.5}$	Nine Months Hiatus	$P(\text{Active})_{.5}$	Six Months Hiatus	$P(\text{Active})_{.5}$
Inactive, correctly classified (%)	84.1569	84.6656	47.8478	.0000	82.6659	87.3881
Active, correctly classified (%)	69.9667	64.3564	89.8708	100.0000	63.5965	53.0703
Overall correctly classified (%)	76.7208	74.0228	82.7467	74.8972	77.1319	77.4289
Inactive but classified active (%)	15.8430	15.3343	52.1521	100.0000	17.3340	12.6120
Active but classified inactive (%)	30.0332	35.6435	10.1291	.0000	36.4034	46.9300
Overall incorrectly classified (%)	23.2791	25.9771	17.2532	25.1072	22.8680	22.5710

Notes: Numbers represent percentage hit rate of the active/inactive class Overall hit rate percentages are weighted according to the distribution of active/inactive customers in the data set.

---

[3] Although it is also possible to optimize for a maximum of correctly classified active or inactive customers depending on the purchase of the marketing action (i.e., reactivation or elimination), we believe that our approach to maximize for the sum of correctly classified active and inactive customers is reasonable because it combines both approaches.

- For $c_{P(Active)} \in \{0, ..., 1\}$, choose $c_{P(Active)}$ so that the sum of correctly classified active and inactive customers is maximized.

For the hiatus heuristic, the algorithm is as follows:

- For $c_{hiatus} \in \{0, ..., \infty\}$, choose c_{hiatus} so that the sum of correctly classified active and inactive customers is maximized.

Table 5 presents an overview of the analysis results. Surprisingly, for all three cohorts, the hiatus heuristic performs slightly better than the more complex Pareto/NBD model. Indeed, the optimal cutoff thresholds for the hiatus heuristic (4 quarters, 40 weeks, and 23 weeks) are close to the managers' and our chosen threshold (3 quarters, 39 weeks, and 26 weeks). If the optimal cutoff thresholds were used instead of expert knowledge, it would result in a marginal gain of only .8302% (airline), .1713% (apparel), and 1.0607% (CDNOW) in terms of overall correctly classified customers.

With respect to the optimal P(Active) thresholds, only for the CDNOW cohort, the optimal value of .44 is close to the natural cutoff value of .5, and its use improves performance by only .17% in terms of overall correctly classified customers. For both the apparel and the airline cohorts, the optimal P(Active) thresholds of .67 and .21, respectively, deviate substantially from the natural cutoff threshold of .5, and the gain from using the optimal threshold rather than the natural cutoff threshold is a considerable 7.9848% and 1.2452%, respectively, in terms of overall correctly classified customers. However, if we carefully examine the analysis results for the apparel cohort, we observe that when the optimal P(Active) of .67 is used, virtually none of the inactive customers are correctly classified (1.009%). The reason is the (estimated) exceptionally long lifetime of an average apparel customer that we already briefly covered in the parameter estimation section. This property causes P(Active) values to be close to 1 for almost all apparel customers and purchase patterns; few customers had P(Active) values in the range of .67–.9. Because we optimized for maximizing the overall correctly classified statistic and given the high percentage of apparel retailer repurchasers, it is more favorable for the optimization algorithm to classify as many active customers correctly as possible. We explore the reasons for these unrealistically high P(Active) values in the "Discussion" section. Nevertheless, the natural cutoff value of .5 may not necessarily be close to its optimal value, as the apparel and airline data sets show. This makes the interpretation of the P(Active) values counterintuitive if a P(Active) value is considered a customer's propensity to repurchase.

Predicting Future Purchase Levels

In this analysis, we focus on the Pareto/NBD and BG/NBD models' capability to predict future purchase levels cumulative for the cohort as a whole and on an individual customer basis.[4] More precisely, we benchmarked both models' predicated number of transactions against a simple management heuristic: Every customer continues to buy at his or her past mean purchase frequency.

We compare the performance of the Pareto/NBD model, the BG/NBD model, and the simple heuristic in predicting cumulated purchases on the basis of the mean absolute percentage error (MAPE) (Leeflang et al. 2000). To measure the performance on the individual customer level, we computed the (root) mean square errors ([R]MSE) for each customer over the predicted and actual transactions in the holdout period (Leeflang et al. 2000). We also computed the mean (R)MSE (median [R]MSE), which represents the mean (median) of all individual customer (R)MSE. Table 6 presents the results of the analysis.

On all three cohorts, the stochastic models outperform the simple heuristic on both the individual and the aggregate levels, and the Pareto/NBD and BG/NBD models perform almost identically. Although the stochastic models deliver decent results on the aggregated level, as the MAPE statistic shows, the results are split on the individual level. The models show poor performance in terms of the mean (R)MSE over all customers, but at least for 50% of the cohorts, the stochastic models predict future purchases precisely, as the median (R)MSE statistics show.

Identifying Future Best Customers

The previous analysis shows that the stochastic models under consideration precisely predict future purchases for

TABLE 5
P(Active) Versus Hiatus Heuristic Using Optimal Thresholds

	Airline		Apparel		CDNOW	
	Hiatus Heuristic	P(Active)	Hiatus Heuristic	P(Active)	Hiatus Heuristic	P(Active)
Optimal cutoff threshold	4 quarters	.21	40 weeks	.67	23 weeks	.44
Inactive, correctly classified (%)	77.109	78.489	46.581	1.009	85.4752	86.3120
Active, correctly classified (%)	77.954	72.343	90.337	99.535	60.3800	56.2863
Overall correctly classified (%)	77.551	75.268	82.918	82.832	78.1926	77.5986
Inactive but classified active (%)	22.891	21.511	53.419	98.991	14.5247	13.6879
Active but classified inactive (%)	22.046	27.657	9.663	.465	39.6199	43.7133
Overall incorrectly classified (%)	22.448	24.731	17.082	17.167	21.8074	22.4014

Notes: Numbers represent percentage hit rate of the active/inactive class. Overall hit rate percentages are weighted according to the distribution of active/inactive customers in the data set.

---

[4] See Schmittlein, Morrison, and Colombo (1987, Equation 22) and Fader, Hardie, and Lee (2005a, Equation 10).

TABLE 6
Summary Statistics for Purchase-Level Prediction

	Airline			Apparel			CDNOW		
Statistic	BG/NBD	Pareto/ NBD	Heuristic	BG/NBD	Pareto/ NBD	Heuristic	BG/NBD	Pareto/ NBD	Heuristic
MAPE	14.582	25.049	28.187	11.3486	12.1242	9.54866	12.15	11.49	55.69
Mean MSE	79.2123	78.2362	95.8921	4.8598	4.8857	5.02916	2.59	2.57	4.89
Median MSE	2.79177	2.3345	7.15625	1.32253	1.3309	1.38375	.061	.027	0
Mean RMSE	4.16536	4.04075	5.24889	1.6338	1.63334	1.63818	.785	.754	1.02
Median RMSE	1.67086	1.5279	2.67512	1.15001	1.15365	1.17633	.248	.166	0

50% of the individuals in a cohort. Nevertheless, this figure tells us little about the applicability of individuals' transaction forecasts in a managerial context. Often, companies implement disproportionate marketing investment strategies on the basis of a customer value rating because it is common for a small percentage of customers to account for a large percentage of revenues and profits (Mulhern 1999). For example, an airline might want to prioritize high-value customers in an overbooking occasion and deny boarding to lower-value customers. Likewise, apparel retailers may want to invite their best customers to special events (i.e., fashion shows), and an online CD store might be interested in sending sample CDs of new albums and/or artists to its best customers.

In this analysis, we assume that a company offers two levels of treatment: "best-customer" treatment and "normal-customer" treatment (Malthouse and Blattberg 2005). Optimally, a customer should receive the best-customer treatment if he or she belongs to the future best customers. Past best customers may not necessarily belong to the group of future best customers (Wangenheim and Lentz 2005). Under the assumption that a customer's future value cannot be estimated perfectly, a company can make two types of classification errors. First, a future best customer may be classified as a future normal customer and thus may be denied the treatment he or she "deserves." This misclassified and, therefore, mistreated customer may spread negative word of mouth or even switch the provider completely. Second, a future normal customer may be misclassified as a future best customer, leading to extra and unjustified spending of scarce marketing resources.

If the complex models under consideration are used to identify future best customers, they need to perform better than a simple management heuristic. This is the subject of our next analysis. More precisely, we try to identify the future 10% (20%, respectively) best customers in the customer bases (in terms of future number of transactions) on the basis of the Pareto/NBD and BG/NBD models' individual customer purchase–level prediction. This classification is benchmarked against yet another simple management judgment rule: The past 10% (20%, respectively) best customers in a customer base will also be the future 10% (20%) best customers.

The results of the analysis appear in Tables 7 and 8. In line with the intention to identify a company's future best customers, the "correctly-classified-as-high" statistic is the one of interest. This statistic represents the fraction of the future best customers who actually have been classified as

TABLE 7
The 10% Best Future Customers

	Airline			Apparel			CDNOW		
Statistic	BG/NBD	Pareto/ NBD	Heuristic	BG/NBD	Pareto/ NBD	Heuristic	BG/NBD	Pareto/ NBD	Heuristic
High, correctly classified (%)	61.09	61.09	57.84	63.15	63.15	70.15	53.92	54.18	61.51
Low, correctly classified (%)	95.22	95.22	94.85	95.63	95.63	94.49	91.08	91.13	86.22
Overall correctly classified (%)	91.76	91.76	90.93	92.18	92.18	91.80	85.06	85.15	82.22
Incorrectly classified high (%)	38.90	38.90	42.15	36.84	36.84	29.84	8.91	8.86	13.77
Incorrectly classified low (%)	4.77	4.77	5.14	4.36	4.36	5.50	46.07	45.81	38.48
Overall incorrectly classified (%)	8.23	8.23	9.06	7.81	7.81	8.19	14.93	14.84	17.77

TABLE 8
The 20% Best Future Customers

	Airline			Apparel			CDNOW		
Statistic	BG/NBD	Pareto/ NBD	Heuristic	BG/NBD	Pareto/ NBD	Heuristic	BG/NBD	Pareto/ NBD	Heuristic
High, correctly classified (%)	64.26	63.40	63.60	67.13	67.13	73.72	61.25	61.69	71.78
Low, correctly classified (%)	89.99	90.60	89.27	91.18	91.18	89.11	84.16	84.33	72.80
Overall correctly classified (%)	84.81	85.12	84.05	86.09	86.09	85.49	77.51	77.76	72.50
Incorrectly classified high (%)	35.73	36.59	36.39	32.86	32.86	26.27	15.83	15.66	27.19
Incorrectly classified low (%)	10.01	9.39	10.72	8.81	8.81	10.88	38.74	38.30	28.21
Overall incorrectly classified (%)	15.18	14.87	15.94	14.37	14.42	15.49	22.48	22.23	27.49

future best customers by the models. The complementary figure is the "incorrectly-classified-as-low" statistic. It represents the fraction of the future best customers who have falsely been classified as future low customers. In four of the six cases, the heuristic performs better that the stochastic models in terms of correctly-classified-as-high customers. Only for the airline data set do the stochastic models outperform the heuristic. We also report the Gini coefficients in Table 9. Instead of focusing only on the 10% and 20% best customers, this measure also includes the models' performance in classifying less valuable customers. The 10% and 20% best-customers statistic and the Gini coefficient provide complementary information; a model can be good at classifying only the 10% or 20% best customers but may be less effective at recognizing the less valuable customers. However, for both the airline and the apparel data sets, the simple heuristic has smaller Gini coefficients than the stochastic models (i.e., it performs better over all customer groups). Because of data constraints, we could not compute the Gini coefficient for the CDNOW data set.

Discussion

Summary of Findings

Many researchers have outlined the usefulness and applicability of the Pareto/NBD model and, more recently, of the BG/NBD model. For example, Krafft (2002), Reinartz and Kumar (2000, 2003), and Wu and Chen (2000) employ the Pareto/NBD model in their work on customer base analysis and CLV prediction. Given this, it is all the more surprising that the current study is the first (1) to validate comprehensively these models' predictions about both the individual customer and the customer base level in noncontractual settings using multiple data sets and (2) to benchmark these models against simple management heuristics that practitioners commonly deploy.

TABLE 9
Gini Coefficients: Best Customer Classification

Data Set	Pareto/NBD	BG/NBD	Heuristic
Airline	.043330	.042874	.041942
Apparel	.071798	.071764	.053581

Notes The CDNOW data set does not provide enough information to compute Gini coefficients.

Recall that by applying the stochastic models, we intended to assist marketing executives in (1) determining active and inactive customers, (2) generating individual customer transaction forecasts to identify the company's future best customers, and (3) determining the future purchase volume of the customer base as a whole. According to our analysis, the applicability of the focal stochastic models seems to be limited to determining the purchase volume of the customer base as a whole. In our analysis, the stochastic models showed superiority over a simple management heuristic. For determining a company's active and inactive customers and for predicting a company's future best customers, the management heuristics we applied worked as well as the stochastic models.

We performed a series of additional analyses, which, for the sake of brevity, cannot be displayed here. Specifically, we varied the holdout period length from 1 to max(holdout period length) (i.e., from 1 to 8 quarters for the airline data set and from 1 to 40 weeks for the apparel data set). At the same time, we held the estimation period constant at 8 quarters and 40 weeks, respectively. Furthermore, we varied the length of the time frame chosen for the apparel data set to have weekly, monthly, and quarterly data available.

The results revealed the same pattern as we reported previously. With regard to active/nonactive classification, the hiatus heuristic classifies an additional 1.5% of the customers correctly compared with the Pareto/NBD model for the apparel data set (averaged over 53 analyses) and an additional 1.34% for the airline data set (averaged over 8 analyses). In terms of identifying the 10% or 20% best customers of the firms, the complex methods are (very slightly) superior in determining only the 10% best customers of the airline. For the best 20% customers and for all analyses related to the apparel data, the hiatus is at least equal and, in most cases, somewhat better than the complex models. (The detailed results of all these analyses, as well as the apparel data set itself, are available on request.)

The sensitivity analysis shows that the length of the holdout period and the aggregation level has little impact on our results. However, it could be argued that this is too little time to classify a customer as inactive. Indeed, depending on their individual purchase rate, customers who have not purchased within the holdout period may well have purchased after that period. Therefore, they are not ultimately inactive (or "dead"), but from a managerial perspective, the lengths of the holdout periods represent a reasonable marketing investment planning horizon. Therefore, it is of hardly any interest to managers whether a customer purchases after the planning horizon. However, even when we restrict our analyses to customers who are active in the holdout period and thus can be certain that our active/inactive assessment is correct, the simple heuristics at least match the performance of the more complex models.

For the apparel data set, we obtained unrealistic lifetime estimates and P(Active) values when we applied the Pareto/NBD model. (A mathematical demonstration and proof that explains why this happens for some data sets are available on request.) In essence, this will be the case when there is a relatively high number of customers in the data set who conducted their last transaction shortly before the end of the observation window and when there are high T values (i.e., an estimation time frame, such that the longer the time frame considered, the higher is the T [weeks, months, quarters]).

Managerial Implications

The finding that P(Active) classifications do not outperform the simple hiatus heuristic in determining active and inactive customers is a devastating result for what has been called the "key result of the NBD/Pareto model" (Reinartz and Kumar 2000, p. 21). As we already mentioned, in at least two cases (the apparel retailer and airline), managers are using the simple hiatus heuristic to determine customer (in)activity in their companies. Their expert assessment coincides almost perfectly with the optimal hiatus length that we determined in our analysis. This is an indication that managerial judgment may well act as a decent estimate of customer (in)activity. Consequently, researchers need to stop recommending the Pareto/NBD model to managers and fellow researchers for this purpose.

For identifying future best customers, the admittedly simple approach of assuming that past best customers are future best customers and the stochastic models deliver unconvincing results, even though we can correct Malthouse and Blattberg's (2005) 20-55 rule to a more positive figure. If we use the Pareto/NBD model or the BG/NBD model, approximately 33% (or less) of the top 20% customers are misclassified, making it a 20-33 rule.[5] Nevertheless, if disproportionate marketing investment decisions are made on the basis of the focal models (i.e., valuable customers receive better service, more perks, and so on, than less valuable customers), these are likely to be inefficient. Scarce marketing resources would be spent on less valuable customers whose behavior does not justify this best-customer treatment. However, many valuable customers who are falsely classified as less valuable customers would not receive the treatment they deserve. Being disenchanted, these customers could switch to a competitor or spread negative word of mouth (Malthouse and Blattberg 2005; Mitchell 2005).

It appears that the managerial applicability of the Pareto/NBD and BG/NBD models is limited to customer equity computation. For this purpose, to both managers and academics, we recommend using the BG/NBD model because of its relatively easier implementation, faster computation, and superior performance compared with the Pareto/NBD model and the simple heuristic. Given the increasing interest in valuing firms on the basis of customer equity (Gupta, Lehman, and Stuart 2004), the stochastic models are good candidates for valuing customer bases in noncontractual settings. Thus, it would be worthwhile to benchmark these models against the approach that Gupta, Lehman, and Stuart (2004) use.

Limitations and Further Research

We determined the top 10% (20%) customers solely on the basis of the number of transactions because the monetary value of transactions was not available in the data. A more managerial top 10% (20%) customer analysis would employ the monetary value of customers. However, we believe that if monetary value were incorporated into our computations, overall model comparison results would not shift substantially; in the end, the Pareto/NBD and BG/NBD models were designed to predict future transactions, not monetary value. Incorporating monetary value would only introduce an additional source of distortion.

For research, much more work still needs to be done. An appealing characteristic of the stochastic models described initially is that they work only on recency and frequency purchase information. The prediction error may partially be explained by the lack of attitudinal information, such as customer satisfaction, repurchase intention, or commitment. However, this does not explain why the stochastic models are not better than models that work on even less and simpler information. Therefore, the question arises whether the information currently used should be substituted by other or augmented by additional information. For example, Wangenheim and Lentz (2005) show that trend in revenues (i.e., the slope of revenue regressed on time) is an important predictor of a customer's life-cycle pattern and improves the accuracy of CLV predictions.

Further research should also address a more general question in predicting future customer purchase patterns: How much purchase information is needed to make predictions about future buyer behavior? For example, in the CDNOW data set (Fader and Hardie 2001), more than 50% of the customers in the data set had not made any purchases since the initial trial (and thus have [0, 0, T] purchase patterns). Can any model, as sophisticated as it may be, make reliable forecasts for a customer who has conducted only one transaction with a supplier? Malthouse and Blattberg (2005) examine the effect of the length of the prediction period on the accuracy of the predictions, and Schmittlein and Peterson (1994) examine how many periods and customers should be included during estimation of the Pareto/NBD model parameters, but to the best of our knowledge, no work has addressed the question of how many transactions a customer needs to have conducted before reliable forecasts can be made.

---

[5] Our analysis confirms Malthouse and Blattberg's (2005) 80-15 rule.

Another notable aspect of our study is that the heuristics the firms used worked astonishingly well. Thus, it would be a worthwhile course for further research to examine how such heuristics emerge in the context of customer management and customer relationship management and how such knowledge can be integrated into relationship management solutions.

Conclusion

This article examines the performance of what have frequently been called state-of-the-art models in customer activity determination and purchase-level prediction in noncontractual settings. To validate these models, we not only used metrics and methods recommended in the statistical literature but also simulated the implementation of those models in managerial practice. However, we find no clear evidence for the superiority of these models for managerially relevant decisions in customer management compared with simple methods that our industry partners used.

As academics, when we lament about practitioners' resistance to using advanced research methods developed in academic research, we too easily forget that model validation in a strict statistical sense is not equivalent to model validation in the spirit of managerial relevance. Although the standard statistical tests are necessary for gaining acceptance in the academic community, they represent a necessary, but not a sufficient, condition for gaining acceptance in the managerial world. Academics must not merely present models' good statistical fit but also suggest how they improve managerial decision making to convince practitioners to adopt them. In the end, this will lead to better cooperation between academics and practitioners. Other examples in the literature on customer purchase predictions show that under certain circumstances, relatively simple heuristics are outperformed by specially developed complex models (Borle, Singh, and Jain 2008). With regard to the Pareto/NBD and BG/NBD models, it is clear that these models carry potential for managerial use in customer management, but it is yet to be shown for which distinct managerial decision they show superior performance.

REFERENCES

Balasubramanian, Sridha, Sunil Gupta, Wagner Kamakura, and Michel Wedel (1998), "Modeling Large Data Sets in Marketing," *Statistica Neerlandica*, 52 (3), 303–324.

Blattberg, Robert C., Gary Getz, and Jacquelyn S. Thomas (2001), *Customer Equity: Building and Managing Relationships as Valuable Assets*. Cambridge, MA: Harvard Business School Press.

Bolton, Ruth N. (1998), "A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction," *Marketing Science*, 17 (1), 45–65.

Bonabeau, Eric (2003), "Don't Trust Your Gut," *Harvard Business Review*, 81 (5), 116–23.

Borle, Sharad, Siddharth Singh, and Dipak Jain (2008), "Customer Lifetime Value Measurement," *Management Science*, 54 (1), 100–112.

Bröder, Arndt (2000), "Assessing the Empirical Validity of the 'Take-the-Best' Heuristic as a Model of Human Probabilistic Inference," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 26 (5), 1332–46.

——— (2003), "Decision Making with the 'Adaptive Toolbox'. Influence of Environmental Structure, Intelligence, and Working Memory Load," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 29 (4), 611–25.

——— and Stefanie Schiffer (2003), "Take the Best Versus Simultaneous Feature Matching: Probabilistic Inferences from Memory and Effects of Representation Format," *Journal of Experimental Psychology: General*, 132 (2), 277–93.

Buchanan, Leigh and Andrew O'Connell (2006), "A Brief History of Decision Making," *Harvard Business Review*, 84 (1), 32–41.

Camerer, Colin F. and Eric J. Johnson (1991), "The Process-Performance Paradox in Expert Judgment: How Can Experts Know So Much and Predict So Badly?" in *Toward a General Theory of Expertise: Prospects and Limits*, K. Anders Ericsson and Jacqui Smith, eds. Cambridge, UK: Cambridge University Press, 195–217.

Creyer, Elizabeth H., James R. Bettman, and John Wayne Payne (1990), "The Impact of Accuracy and Effort Feedback and Goals on Adaptive Decision Behavior," *Journal of Behavioral Decision Making*, 3 (1), 1–16.

Donkers, Bas, Peter C. Verhoef, and Martijn de Jong (2007), "Modeling CLV: A Test of Competing Models in the Insurance Industry," *Quantitative Marketing and Economics*, 5 (2), 163–90.

Ehrenberg, Andrew S.C. (1988), *Repeat-Buying, Theory and Applications*, 2d ed. London: Griffin.

Fader, Peter S. and Bruce G.S. Hardie (2001), "Forecasting Repeat Sales at CDNOW: A Case Study," *Interfaces*, 31 (Part 2 of 2), 94–107.

———, ———, and Ka Lok Lee (2005a), "'Counting Your Customers the Easy Way: An Alternative to the Pareto/NBD Model," *Marketing Science*, 24 (2), 275–85.

———, ———, and ——— (2005b), "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42 (November), 415–30.

Gigerenzer, Gerd, Peter M. Todd, and The ABC Research Group (1999), *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Goldberg, Lewis R. (1968), "Simple Models or Simple Processes? Some Research on Clinical Judgments," *American Psychologist*, 23 (7), 483–96.

Gupta, Sunil, Donald R. Lehmann, and Jennifer A. Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41 (February), 7–18.

Hayashi, Alden M. (2001), "When to Trust Your Gut," *Harvard Business Review*, 79 (2), 59–65.

Helsen, Kristiaan and David C. Schmittlein (1993), "Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models," *Marketing Science*, 12 (4), 395–414.

Ho, Teck-Hua, Young-Hoon Park, and Yong-Pin Zhou (2006), "Incorporating Satisfaction into Customer Value Analysis:

Optimal Investment in Lifetime Value," *Marketing Science*, 25 (3), 260–77.

Hughes, Arthur M. (2006), *Strategic Database Marketing*, 3d ed. New York: McGraw-Hill.

Jain, Dipak and Siddhartha S. Singh (2002), "Customer Lifetime Value Research in Marketing: A Review and Future Directions," *Journal of Interactive Marketing*, 16 (2), 34–47.

Kamakura, Wagner, Carl. F. Mela, Asim Ansari, Anand Bodapati, Pete Fader, Raghuram Iyengar, Prasad Naik, Scott Neslin, Baohong Sun, Peter C. Verhoef, Michel Wedel, and Ron Wilcox (2005), "Choice Models and Customer Relationship Management," *Marketing Letters*, 16 (3–4), 279–91.

Krafft, Manfred (2002), *Kundenbindung und Kundenwert*. Heidelberg: Physica-Verlag.

Lee, Michael D. and Tarrant D.R. Cummins (2004), "Evidence Accumulation in Decision Making: Unifying the 'Take the Best' and the 'Rational' Models," *Psychonomic Bulletin & Review*, 11 (2), 343–52.

Leeflang, Peter, Dick R. Wittink, Michel Wedel, and Philippe A. Naert (2000), *Building Models for Marketing Decisions*. Boston: Kluwer Academic.

Li, Shaomin (1995), "Survival Analysis," *Marketing Research*, 7 (Fall), 17–23.

Malthouse, Edward and Robert Blattberg (2005), "Can We Predict Customer Lifetime Value?" *Journal of Interactive Marketing*, 19 (1), 2–16.

Mitchell, Alan (2005), "Who Knows the Relative Worth of Customers?" *Precision Marketing*, 17 (49), 12.

Morwitz, Vicki G. and David C. Schmittlein (1998), "Testing New Direct Marketing Offerings: The Interplay of Management Judgment and Statistical Models," *Management Science*, 44 (5), 610–28.

Mulhern, Frank (1999), "Customer Profitability Analysis: Measurement, Concentration, and Research Directions," *Journal of Interactive Marketing*, 13 (1), 25–40.

Newell, Ben R., Tim Rakow, Nicola J. Weston, and David R. Shanks (2004), "Search Strategies in Decision Making: The Success of 'Success,'" *Journal of Behavioral Decision Making*, 17 (2), 117–37.

——— and David R. Shanks (2003), "Take the Best or Look at the Rest? Factors Influencing 'One-Reason' Decision Making," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 29 (1), 53–65.

———, Nicola J. Weston, and David R. Shanks (2003), "Empirical Tests of a Fast-and-Frugal Heuristic: Not Everyone 'Takes-the-Best,'" *Organizational Behavior & Human Decision Processes*, 91 (1), 82–96.

Novo, Jim (2004), *Drilling Down: Turning Customer Data into Profits with a Spreadsheet*, 2d ed. St. Petersburg, FL: The Drilling Down Bookstore.

Payne, John Wayne, James R. Bettman, and Eric J. Johnson (1993), *The Adaptive Decision Maker*. Cambridge, UK: Cambridge University Press.

Reinartz, Werner J. and V. Kumar (2000), "On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implications for Marketing," *Journal of Marketing*, 64 (October), 17–35.

——— and ——— (2003), "The Impact of Customer Relationship Characteristics on Profitable Lifetime Duration," *Journal of Marketing*, 67 (January), 77–99.

Rieskamp, Jörg (2006), "Perspectives of Probabilistic Inferences: Reinforcement Learning and an Adaptive Network Compared," *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 32 (6), 1355–70.

——— and Ulrich Hoffrage (1999), "When Do People Use Simple Heuristics and How Can We Tell?" in *Simple Heuristics That Make Us Smart*, Gerd Gigerenzer, Peter M. Todd, and The ABC Research Group, eds. New York: Oxford University Press, 141–68.

——— and ——— (2008), "Inference Under Time Pressure: How Opportunity Costs Affect Strategy Selection," *Acta Psychologica*, forthcoming.

——— and Phillip E. Otto (2006), "SSL: A Theory of How People Learn to Select Strategies," *Journal of Experimental Psychology: General*, 135 (May), 207–236.

Roel, Raymond (1988), "Direct Marketing's 50 Big Ideas," *Direct Marketing*, 50 (May), 45–52.

Rust, Roland and Tuck S. Chung (2006), "Marketing Models of Service and Relationships," *Marketing Science*, 25 (6), 560–80.

Schmittlein, David C., Donald G. Morrison, and Richard Colombo (1987), "Counting Your Customers: Who Are They and What Will They Do Next?" *Management Science*, 33 (1), 1–24.

——— and Robert A. Peterson (1994), "Customer Base Analysis: An Industrial Purchase Process Application," *Marketing Science*, 13 (1), 41–67.

Sharma, Subhash (1996), *Applied Multivariate Techniques*. New York: John Wiley & Sons.

Verhoef, Peter C., Penny N. Spring, Janny C. Hoekstra, and Peter S.H. Leeflang (2002), "The Commercial Use of Segmentation and Predictive Modeling Techniques for Database Marketing in the Netherlands," *Decision Support Systems*, 34 (4), 471–81.

Wade, Betsy (1988), "Mileage Points Can Fly Away," *The New York Times*, (November 20), (accessed January 14, 2008), [available at http://query.nytimes.com/gst/fullpage.html?res=940DE1DC1F3DF933A15752C1A96E948260].

Wangenheim, Florian and Patrick Lentz (2005), "Customer Portfolio Analysis: Applying Financial Risk and Volatility Measures to Customer Segmentation and Risk-Adjusted Lifetime Value Determination," (October), (accessed August 17, 2007), [available at http://ssrn.com/abstract=782064].

Wu, Couchen and Hsiu-Li Chen (2000), "Counting Your Customers: Compounding Customer's In-Store Decisions, Interpurchase Time and Repurchasing Behavior," *European Journal of Operational Research*, 127 (1), 109–119.

Khaled Aboulnasr, Om Narasimhan, Edward Blair, & Rajesh Chandy

Competitive Response to Radical Product Innovations

Radical product innovations are often agents of creative destruction. They threaten to destroy existing market positions, and yet they often yield vast new market opportunities. This article examines how competitors respond to the introduction of radical product innovations. The authors argue that competitive response to radical product innovations is inherently different from response to the incremental innovations that are typically studied in existing research. They introduce the dual concepts of market expansion and entry thresholds to develop new hypotheses about competitive response. Some of these hypotheses contradict prior literature. Using objective data from the U.S. pharmaceutical industry between 1997 and 2001, they estimate a shared-frailty hazard model to explain the competitive response to radical product innovations. The results show that the likelihood of competitive response is substantially higher when the introducing firm is large or market dependent. Moreover, the response is highest when the innovation is introduced in a small market by a large firm. These results contradict those from much prior research on competitive response to product innovation.

Keywords: innovation, competition, market expansion, entry thresholds

The specter of competition looms large in all product introductions. Introducers try to predict which competitors will respond and when. Some competitors scramble to introduce products of their own. Others refrain from action, perhaps from a fear of retaliation, lack of financial resources, sloth induced by inertia, or a fear of cannibalizing existing products (Kuester, Homburg, and Robertson 1999; Rhoades 1973; Tellis and Golder 2001).

Given the central role of competition in the economic system, the study of competitive response is essential for any understanding of business actions. Managers need to incorporate competitive response into their financial projections as they decide how much to invest in new products, or their dreams of riches could easily turn into dust. Researchers need to ensure that the reaction functions in their models of competitive interaction are accurate, or their insights could mislead (Bowman and Gatignon 1995; Moorthy 1985; Weitz 1985). Policy makers need to be able to predict competitive response, or their interventions could be unwise.

This study examines competitive response to radical product innovations. Radical product innovations differ from other new products in that they have substantially different technology and substantially higher benefits than existing products (e.g., Chandy and Tellis 1998). As such, radical product innovations are riskier than other product introductions and demand more resources (Sorescu, Chandy, and Prabhu 2003). They are also more likely to destabilize markets and cause customers to reconsider existing purchase patterns. Thus, they not only threaten existing competitive positions but also offer new market opportunities. However, the growing literature on radical product innovation is largely silent on the issue of competitive response, and the large and established literature on competitive response to product introductions often overlooks the introduction of radical product innovations (see Bowman and Gatignon 1995; Kuester, Homburg, and Robertson 1999; Min, Kalwani, and Robinson 2006; Robinson 1988; Sheremata 2004).

In this article, we extend and connect these two important streams of research in marketing strategy—on competitive response and radical product innovation—and add to the literature both theoretically and empirically. From a theoretical perspective, we use signaling arguments (e.g., Heil and Langvardt 1994; Heil and Robertson 1991; Porter 1980; Prabhu and Stewart 2001; Robertson, Eliashberg, and Rymon 1995) and introduce the dual concepts of market expansion and entry thresholds to develop new hypotheses about competitive response. In our terms, market expansion refers to an increase in the size of a market, and an entry threshold for any given firm refers to the minimum size of a potential market that would prompt the firm to introduce a new product into the market.[1] To the best of our knowledge, no prior research has used these constructs to explain competitive response. We attempt to build a "home-grown" theory of the type that Rust (2006, pp. 1–2) calls for to

Khaled Aboulnasr is Assistant Professor of Marketing, Lutgert College of Business, Florida Gulf Coast University (e-mail: kaboulna@fgcu.edu). Om Narasimhan is Associate Professor of Marketing (e-mail: naras002@umn.edu), and Rajesh Chandy is James D. Watkins Chair in Marketing (e-mail: rchandy@umn.edu), Carlson School of Management, University of Minnesota. Edward Blair is Professor of Marketing, C.T. Bauer College of Business, University of Houston (e-mail: blair@uh.edu). The authors thank Akshay Rao for valuable suggestions on a previous draft of this article.

---

[1] We define entry thresholds at the firm level, unlike authors such as Bresnahan and Reiss (1991), who define entry thresholds at the market level. In addition, by entry thresholds, we do not imply entry by a firm into a new market; we are interested in the introduction of a new product into a market regardless of whether the firm already operates in that market (Chandy and Tellis 1998).

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (May 2008), 94–110

explore previously unexplored issues in competitive response.

Some of the hypotheses we develop contradict prior literature, which has argued that competitors are less likely to respond to the introduction of new products by larger firms than to that of smaller firms because of a fear of retaliatory behavior (Bowman and Gatignon 1995; Shankar 1999). However, in the context of radical product innovations, we argue that competitors are more likely to respond to larger than to smaller firms. Indeed, we find that radical product introductions by large firms are almost twice as likely to provoke a reaction as those by small firms. Prior literature has also argued that competitors are less likely to respond to product introductions in smaller than in larger markets (Gruca, Kumar, and Sudharshan 1992; Gruca and Sudharshan 1995; Shankar 1999). In contrast, we find that competitors are more likely to respond to radical product innovations in smaller than in larger markets if the firm introducing the innovation is large. In addition, we examine how response to radical product innovations is influenced by market dependence, a variable largely overlooked in prior research (see Montaguti, Kuester, and Robertson 2002). We argue that competitors are more likely to respond to radical product innovations by firms that are more market dependent (i.e., firms that derive a larger part of their revenues from that particular market). We find that an increase of one standard deviation in the market dependence of the firm introducing the radical innovation makes competitor reaction almost 50% more likely.

From an empirical perspective, we address several methodological limitations that have been cited in previous research on competitive response. One such limitation is small sample size (Kuester, Homburg, and Robertson 1999; Robinson 1988). For example, Yip (1982) uses only 37 observations, and Shankar (1999) uses 23 new product entries and 59 response observations. A second limitation is completeness of the data—namely, measuring response for only a limited number of competitors (Robinson 1988; Shankar 1997). A third limitation is the possibility of self-reporting bias in questionnaire measures (Kuester, Homburg, and Robertson 1999; Robinson 1988). By focusing on the pharmaceutical industry and using data from multiple sources, we attempt to address each of these issues. The pharmaceutical industry has been popular in studies of innovation because it is a multibillion dollar industry that is driven by innovation. Within this important industry, we are able to obtain objective measures of whether new products constitute radical product innovations (Sorescu, Chandy, and Prabhu 2003), as well as objective measures of which competitors responded and when they responded. We develop a comprehensive database of innovation and response over a five-year period, covering more than 50 radical product innovations and more than 700 observations of competitive response, the largest sample obtained to date for this type of research. Our database enables us to develop a richly specified empirical model that accounts for all variables relevant to the testing of our hypotheses, as well as other variables that have been suggested in prior research and variables that address unique features of our empirical context.

The focus of this article is on competitive responses in the form of product introductions. Although competitors can also respond through other elements of the marketing mix, we focus on product responses because prior research has suggested that responses to competitive actions tend to be reciprocal (e.g., product responses for product actions, price responses for price actions; see Axelrod 2002; Bowman and Gatignon 1995). In addition, Kuester, Homburg, and Robertson (1999) argue that response is especially likely to be on the product dimension when it is prompted by the introduction of highly innovative products.

We organize the rest of this article as follows: In the next section, we present a conceptual framework and develop our hypotheses. Following this, we describe our research methods, including variable operationalizations, data sources, and analysis procedures. We then present the results of the analysis. Finally, we conclude by identifying possible limitations and discussing several implications for scholarship and practice.

Conceptual Framework

We focus on competitive response to radical product innovations: specifically, on the likelihood that competitors will respond to an innovation by introducing products of their own. A radical product innovation is a new product that uses significantly different technology and offers significantly greater customer benefits per dollar than existing products (Chandy and Tellis 1998). In many ways, radical product innovations are the "home runs" of product innovation and have the potential to be extremely lucrative. For example, in the pharmaceutical industry, Sorescu, Chandy, and Prabhu (2003) show that stock market returns to the introduction of radical product innovations can be in the billions of dollars. This said, the effects of radical product innovations are not uniformly positive or straightforward. Such innovations have the potential for three important effects as they relate to existing markets (e.g., Chandy and Tellis 1998; Chen and Miller 1994; Christensen 1997): (1) market expansion, (2) cannibalization, and (3) destabilization.

From the perspective of the market as a whole, radical product innovations imply a high potential for market expansion. This is perhaps the most striking difference in outcomes between radical and incremental innovations (Montaguti, Kuester, and Robertson 2002; see also Mahajan, Sharma, and Buzzell 1993). In general, the impact of incremental innovations is to redistribute shares within an existing market. Conversely, radical product innovations provide significantly greater benefits than were previously available and thus may substantially increase the size of the market (Golder and Tellis 1997; Sorescu, Chandy, and Prabhu 2003). Consider the following examples from the U.S. pharmaceutical industry (IMS Health Global Services 2004; all dollar values are inflation adjusted and listed in 1998 dollars):

- The size of the bowel syndrome category was $377 million in 1999, the year GlaxoSmithKline introduced Lotronex. The size of this category became $1.14 billion by 2003.

- The size of the sexual dysfunction category was $587 million in 1998, the year Pfizer introduced Viagra. The size of the category became $1.2 billion by 2003.
- The size of the arthritis category was $1.15 billion in 1998, the year Merck introduced Celebrex. The size of the category became $8.30 billion by 2003.

From the perspective of innovating firms, radical product innovations can result in substantial cannibalization of existing business (Chandy and Tellis 1998; Govindarajan and Kopalle 2004). An element of cannibalization is sales cannibalization, whereby innovations take away sales from the firm's existing products in the category. Another element is the cannibalization of specialized investments, whereby innovations reduce the value of investments that are tied to existing products (Nijssen, Hillebrand, and Vermeulen 2005). Innovating firms must incorporate the potential for cannibalization in their decision making leading up to the introduction of an innovation (Kerin, Harvey, and Rothe 1978).

From the perspective of incumbent competitors that already have products within the category of an innovation, radical product innovations imply a high potential for market destabilization (Schumpeter 1942). By redefining the product category's benefit space, a radical product innovation may not only seize business from existing competitors but also reposition existing products relative to each other (Van Heerde, Mela, and Manchanda 2004). For example, a product that previously held a distinctive performance position may collapse into a generic "old generation" position, placing it closer to products from which it was previously well differentiated, and may require a "new-and-improved" model to regain its previous position. The potential for destabilization makes competitive response much more likely for radical product innovations than for many other competitive actions (Chen and Miller 1994).

Market Expansion, Entry Thresholds, and Competitive Response

What factors influence the likelihood of competitors' response to a radical product innovation? Various factors might be argued on the basis of the general literature on competitive response (see Bowman and Gatignon 1995; Chen and Miller 1994; Kuester, Homburg, and Robertson 1999; Montaguti, Kuester, and Robertson 2002; Robertson, Eliashberg, and Rymon 1995; Robinson 1988). However, with specific respect to radical product innovations, our research focuses on the potential for market expansion and the role of entry thresholds in signaling the extent of market expansion expected. Whatever the other effects of a radical product innovation may be, we argue that competitors are more likely to respond by introducing their own products when some aspect of the radical product innovation provides them with signals that it will increase the size of the market.[2]

What factors have these effects? Some of the most credible signals in this context are related to the nature of firms that introduce innovations (Prabhu and Stewart 2001). The very act of product introduction reveals information about the introducer's expectations about the potential of the market it is entering. The introducer reveals this information by virtue of (1) who it is and (2) which market it enters. Incumbent competitors incorporate this information into their own decision calculus as they determine when, whether, and how to respond.

We argue that the likelihood of response is greater when potential respondents observe product introduction by firms that have higher entry thresholds (i.e., firms that would only introduce a product if the market has high potential). Moreover, the impact of this signal is the greatest when such firms introduce products in markets that were previously viewed as having low potential.

What is the profile of a firm with high entry thresholds? Although factors unique to each firm are likely to play a role, we argue that two factors systematically signal a firm's entry threshold: (1) the firm's size and (2) its dependence on the market it is entering. Our focus on firm size and market dependence in the context of radical product innovation is consistent with the literature on radical product innovation, which also emphasizes the importance of these two factors (see Chandy, Prabhu, and Antia 2003; Chandy and Tellis 2000).

First, in terms of firm size, larger firms tend to have higher entry thresholds. This is in part because small markets do not meet the growth needs of large firms (Christensen 1997). As firms become larger, their reference points for what constitutes attractive markets also become larger. The prospect of a $10 million business might cause great excitement in a small firm, but this might be met with a shrug in many large firms. As such, the introduction of a radical product innovation by a larger firm is more likely to convey an expectation of market expansion, especially if the current market is relatively small. When confronted with a large firm entering a small market, potential responders are likely to ascribe this otherwise atypical behavior to an expectation of market expansion; that is, the large firm expects that the market will expand substantially as a result of the introduction of the radical product innovation.

Second, in terms of market dependence, a market-dependent firm derives large parts of its revenue from that particular market. Firms that are highly dependent on a market also tend to have high entry thresholds. This is because by introducing a radical product innovation in the market, the firm is likely to cannibalize the sales of its existing products (Chandy and Tellis 1998). Therefore, the radical product innovation faces a greater burden of expectations in a market-dependent firm than in other firms (Kerin, Harvey, and Rothe 1978). Firms with higher levels of market dependence are most likely to introduce a radical product innovation if they expect enough market expansion to compensate for the cannibalization of existing products.

---

[2] Note that we do not use the term "signal" as does Spence (1974), who argues that a sender uses signals strategically. In line with a substantial literature in marketing (e.g., Heil and Langvardt 1994; Heil and Robertson 1991; Robertson, Eliashberg, and Rymon 1995) and elsewhere (e.g., Porter 1980), we define a signal as an action by a firm "that conveys information about its intentions and abilities" (Prabhu and Stewart 2001, p. 63).

Hypotheses

Effect of Introducer Size

The existing literature on competitive response to product introductions argues that incumbent competitors are less likely to respond to the introduction of new products by larger than by smaller firms because of the deterrent effect of larger resources (Bowman and Gatignon 1995; Shankar 1999). For example, Shankar (1999) notes that it may be unwise to respond to large-scale entrants because of the fear of a war of attrition. Similarly, Bowman and Gatignon (1995) argue that competitors are less likely to respond to a new product introduced by a strong firm with large resources because of fear of retaliation. However, in the context of radical product innovation, we argue the opposite. Because radical product innovations are inherently destabilizing (Schumpeter 1942), incumbent firms are much less likely to hold back in an effort to limit competition. Instead, a primary factor in response will be incumbents' assessments of the extent to which the innovation will be successful and the extent to which it will expand the market, as we discussed previously. Incumbents will surmise that a larger firm, given its high entry threshold, will introduce an innovation only if it expects the market potential for the innovation to be large. As such, we posit that competitive response to a radical product innovation will be more likely when the introducer is a large firm than when it is a small firm because the entry of the former is more likely to signal expected market expansion. In summary, we hypothesize the following:

> H_1: The greater the size of the firm introducing a radical product innovation, the greater is the likelihood of competitive response.

Effect of Introducer Market Dependence

We also argue that there is a greater likelihood of response to radical product innovations introduced by firms with higher levels of market dependence. Market-dependent firms introducing a radical product innovation have more to lose by disrupting the market (Heide and Weiss 1995; Montaguti, Kuester, and Robertson 2002). The very notion of "more to lose" suggests that the entry threshold for such firms is high. For a market-dependent firm to introduce a radical product innovation, the new product must promise returns that are large enough to compensate for the likely loss of sales and investments associated with existing products on which the firm depends. The resultant implication of possible market expansion is likely to stimulate competitive response. Thus:

> H_2: The greater the market dependence of the firm introducing a radical product innovation, the greater is the likelihood of competitive response.

Interaction Effects of Market Size

The information on potential for market expansion that is conveyed by the size and market dependence of the introducing firm is likely to be especially powerful when the size of the market in which the radical product innovation is introduced is small. In such a case, the information inherent in the decision of a large or market-dependent firm to enter the market takes on even greater significance. The reasoning is that the large firm might detect significant potential in the market that has previously not been realized (Christensen 1997). This will motivate competitors to respond. An identical argument applies to the case of market dependence. Thus:

> H_3: The likelihood of competitive response is greater when (a) a large firm introduces a radical product innovation in a small market than when it does so in a large market and (b) a market-dependent firm introduces a radical product innovation in a small market than when it does so in a large market.

Other Variables of Interest

Our hypotheses focus on factors that signal the potential for market expansion, given the entry thresholds of firms that introduce radical product innovations. However, these are not the only variables that might influence competitive response. Other variables that speak to the attractiveness of the market, or the motivation or capabilities of incumbent competitors, may also have an influence. Most of these variables have already been covered in the literature on competitive response. To avoid repetition, we do not state hypotheses for these variables, but we include them as control variables in our analyses and note their likely effects.

Prior literature has considered the market growth rate an indicator of market attractiveness and, thus, an antecedent of competitive response (Bowman and Gatignon 1995; Kuester, Homburg, and Robertson 1999; Shankar 1999). Empirical results for this variable have been mixed (Shankar 1999). One scenario is that competitive response will be stronger in growing markets because competitors view the market as highly attractive and will fight for it. Bowman and Gatignon (1995) and Kuester, Homburg, and Robertson (1999) support this view. An alternative is that competitive response will be weaker in growing markets, perhaps because market growth is already taxing the resources of competitors (Tellis and Golder 2001) or because competitors are already satisfied with their performance (Bowman and Gatignon 1995).

Another market-related factor that has been studied in the literature is market concentration. Thus, it could be argued that competitive response is more likely in concentrated markets, partly because competitors can monitor their rivals more carefully (Blundell, Griffith, and Van Reenen 1999; Montaguti, Kuester, and Robertson 2002; Nickell 1996). Concentration may also be related to market attractiveness, in that high concentration could indicate that the industry has found only a few solutions to a customer problem, each of which may have limitations, leaving a higher upside for new solutions. To the extent that higher concentration indicates higher upside in the market, there should be a greater likelihood of competitive response. Conversely, some research indicates that competitive response may be less likely in concentrated markets, perhaps because of mutual forbearance (Bernheim and Whinston 1998; Chen and MacMillan 1992).

With respect to market attractiveness, another relevant variable is current market size. Apart from the interaction effects that we hypothesized, it can be argued that larger markets will motivate stronger competitive response to radical product innovations because competitors are more willing to fight for the market. Alternatively, there could be a lower likelihood of competitive response in larger markets, perhaps because competitors are more likely to be inertia prone in markets in which sales are plenty (Chandy, Prabhu, and Antia 2003).

A competitor-related variable relevant to competitive response is the size of the competitor firm. Prior research has argued that larger competitors are less likely to respond to marketing actions because of bureaucratic inflexibility or inertia (Tornatzky and Fleischer 1990). Conversely, it can be argued that larger firms have greater resources and, thus, greater capabilities for response. In the context of radical product innovations, in which the destabilizing potential of the innovation may provide an inherent motivation to respond, these capabilities may result in larger firms being more likely to respond.

Another competitor-related variable relevant to competitive response is the market dependence of the competitor firm. As with the introducing firm, a competitor's market dependence indicates the importance of the market to that firm. Market-dependent competitors have a greater stake in the market, which may motivate a stronger defensive posture and a greater likelihood of response (Chen and MacMillan 1992).

Finally, two other factors might potentially affect competitive response. First, competitors might be less likely to respond if they have recently introduced a new product in the same category because of either commitment to that product or depletion of product development resources. Second, there may be order-of-response effects; specifically, later firms may be less likely to respond because earlier competitors have claimed preemptive market positions, or alternatively, later firms may be more likely to respond because of a bandwagon effect.

The following section discusses how we test each of our arguments. Specifically, it provides details regarding (1) our considerations in choosing an empirical context, (2) operationalizations and data sources for radical product innovation and competitive response, (3) operationalizations and data sources for other variables, and (4) our model specifications. Figure 1 summarizes the conceptual framework.

Method

We test our hypotheses in the context of the pharmaceutical industry. The general method is as follows: We identify all radical product innovations in the pharmaceutical industry over a five-year period, from 1997 through 2001; there are 52 such innovations. In the three years following each innovation, we track whether each competitor in the product category responded by introducing its own new product and, if so, when. This gives us 714 total observations of whether (and when) competitors responded to radical product innovations. We analyze competitive response with a hazard model formulation that models the likelihood of any given competitor responding in any given period as a function of various independent variables—those for which we state hypotheses as well as various controls. The goal of a hazard model is to examine longitudinal and cross-sectional effects in duration times and to give probabilistic or expectation-based predictions (DuWors and Haines 1990; Grimshaw et al. 2005).

Empirical Context

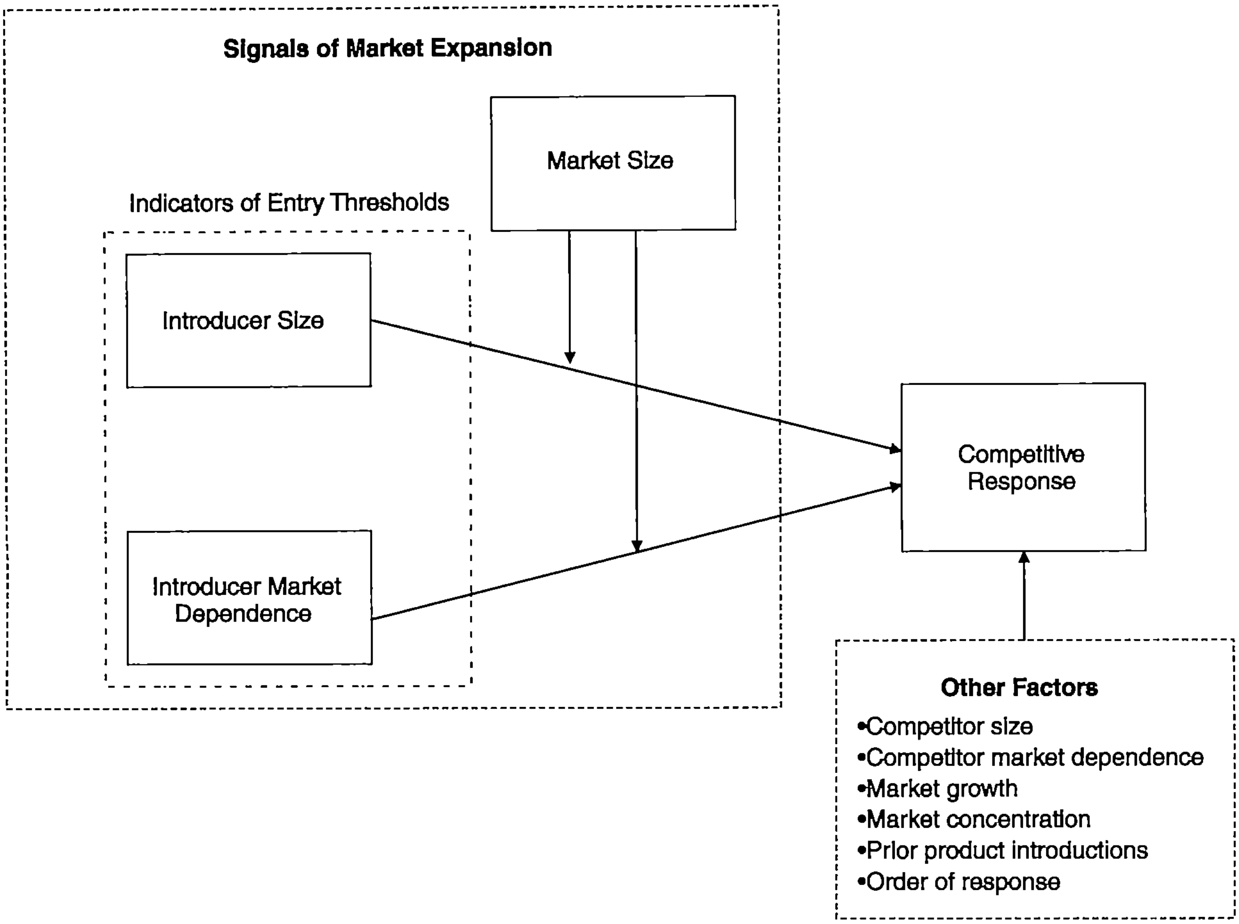
The model presented in this study describes the likelihood of competitive response to radical product innovations as a function of introducer firm size, introducer market dependence, the interaction of these two variables with market size, and various control variables. To test our arguments, we need an empirical context that has many product categories and considerable variety in firm size and market size. It also should be a context in which numerous introductions of radical product innovations can be objectively identified. The pharmaceutical industry fits all our requirements.

First, the pharmaceutical industry is a huge industry, covering a large number of therapeutic categories with substantial variation in our variables of interest. Second, the pharmaceutical industry offers a substantial number of radical product innovations, which are a driving force behind the growth of the industry (Scherer 2000). Third, the pharmaceutical industry provides an objective, unbiased measure of radical product innovation, as we discuss subsequently.

Operationalizations and Data Sources for Radical Product Innovation and Competitive Response

The U.S. Food and Drug Administration (FDA) classifies new drugs on two dimensions: chemical composition and therapeutic potential (for details, see Sorescu, Chandy, and Prabhu 2003). Regarding chemical composition, the FDA classifies a drug as a "new molecular entity" versus a composition that is a new formulation, new combination, or new usage of existing chemistry. New molecular entities represent the most technologically advanced products, with an active ingredient that has never been in the market before. Regarding therapeutic potential, a new drug that represents significant therapeutic benefits compared with all existing drugs is given "priority review" by the FDA, whereas a new drug that has healing features similar to drugs already in the market is given a standard review. These FDA classification dimensions—chemical composition and therapeutic potential—correspond exactly to the technology and benefit dimensions that Chandy and Tellis (1998) and Sorescu, Chandy, and Prabhu (2003) use to define radical product innovations. Thus, we classify a new product introduction as a radical product innovation if it meets the criteria of being a new molecular entity (a distinct advancement in technology) and receiving priority review (a distinct advancement in customer benefits). This operationalization

FIGURE 1
A Conceptual Model of Competitive Response to Radical Product Innovation


Signals of Market Expansion
Market Size
Indicators of Entry Thresholds
Introducer Size
Introducer Market Dependence
Competitive Response
Other Factors
•Competitor size
•Competitor market dependence
•Market growth
•Market concentration
•Prior product introductions
•Order of response

of radical product innovation is identical to that used in previous research on this topic in marketing (Sorescu, Chandy, and Prabhu 2003; Wuyts, Stremersch, and Dutta 2004) and elsewhere (Yeoh and Roth 1999).

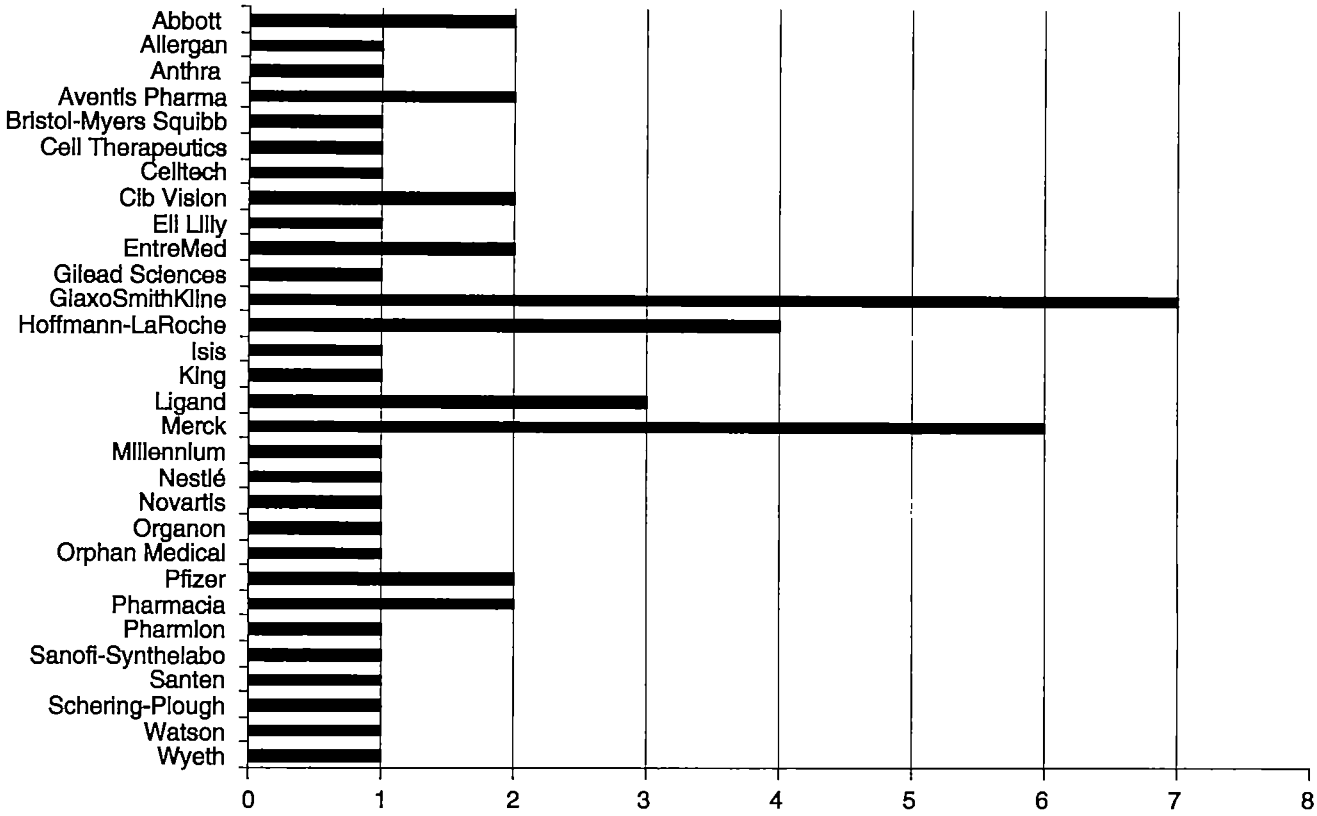
Our data contain 52 radical product innovations introduced by 32 different companies in 27 therapeutic categories over the 1997–2001 period. On average, each innovation faced 14 incumbent products within its category (high = 47, low = 2), for a total of 714 observations for potential response by incumbents. Figure 2 shows the number of radical product innovations introduced by each company, and Figure 3 shows the number of radical product innovations per therapeutic category. This information was acquired from the NDA Pipeline. The NDA Pipeline is a comprehensive database that tracks the FDA approval process for new drugs through all phases of development until product introduction.

We measured response over the three years following the introduction of a radical product innovation. For example, if a radical product innovation was introduced in March 2001, we measured response through March 2004. Previous research has measured response over windows ranging from six months to two years (Bowman and Gatignon 1995; Kuester, Homburg, and Robertson 1999; Robinson 1988; Shankar 1999). We used three years because it is twice the average period of one and a half years that competitors took to respond to breakthrough pharmaceutical products in the 1990s (Pharmaceutical Research and Manufacturers of America 2003). A three-year period allows us to examine responses without causing substantial censoring in the data.

Within the three-year observation window, we measured each competitor's response as the number of days after the introduction of the innovation before the competitor introduced a product of its own. We treated competitors that did not respond by the last date in the window as censored observations in the hazard model. Overall, approximately half the competitors introduced products within the three-year window. We also obtained data on competitive product introductions through the NDA Pipeline.

Certain points might be noted about the process of innovation and response in this industry. First, as in most industries, competitors have some awareness of an innovation before it is actually introduced to the market. In most industries, this type of information is available through sources such as industry gossip, personnel movement, trade publications, and company announcements. In pharmaceuticals, this type of information is available through the regulatory process. New drugs in the United States go through a multi-

FIGURE 2
Number of Radical Product Innovations per Introducing Firm



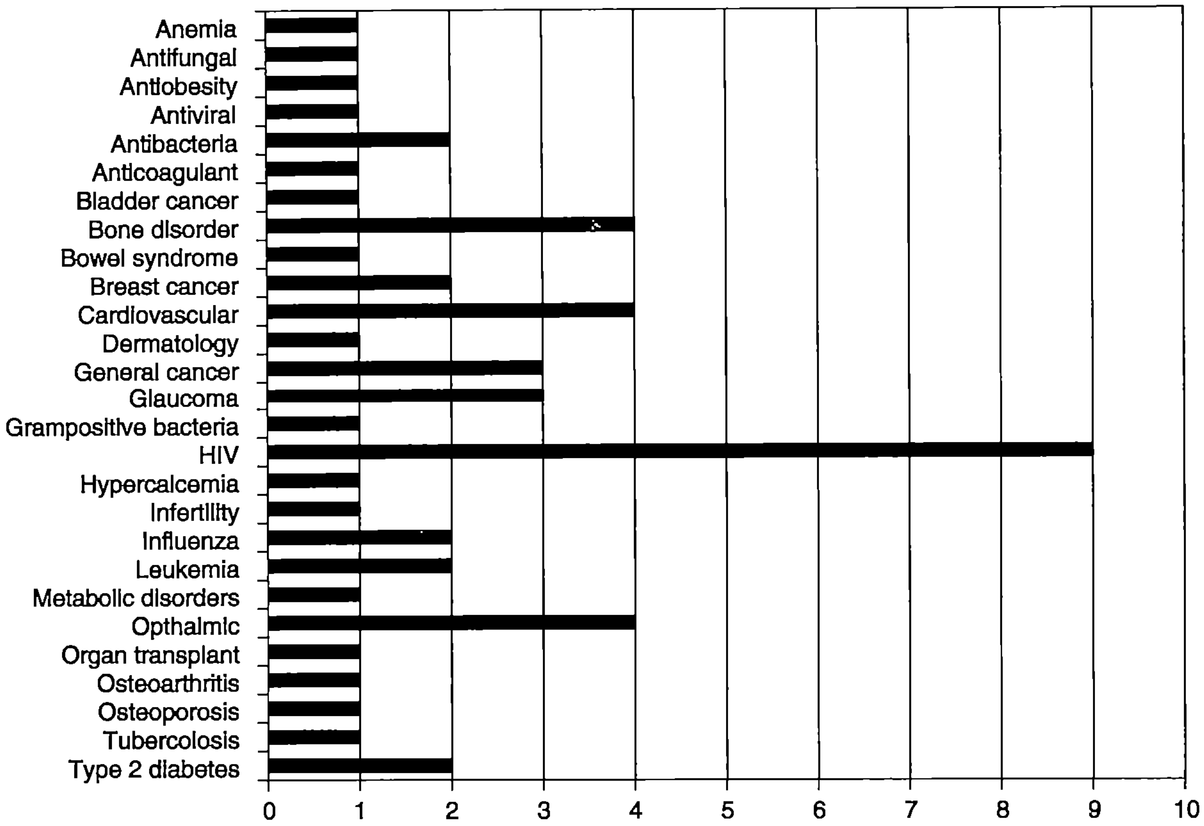
stage FDA approval process (see Figure 4; FDA 1999). The drug sponsor first applies for an investigational new drug application based on preclinical animal tests. If approved, the drug proceeds to three phases of clinical trials with humans. If the drug is approved through all three phases of clinical trials, the drug sponsor files a new drug application, and if this is approved, the sponsor can introduce the drug. Records of approvals at each stage are public information, so competitors know that a new drug may be coming before it is introduced.

Second, as in most industries, a company may be working on an innovation, but this does not necessarily mean that a successful product will result. This may be particularly true in the pharmaceutical industry. Only 1 in 1000 new drug compounds makes it as far as the clinical studies phase, and according to the Tufts Center for the Study of Drug Development, only 1 in 5 drugs that make it to clinical testing is ultimately approved for introduction (www.fda.gov). Therefore, the various stages of preintroductory product development have uncertain implications, and competitors that respond to products in the early stages are often chasing ghosts.

Third, radical product innovations in the pharmaceutical industry are typically protected by patents. This means that competitors cannot easily respond by duplicating the innovation; rather, they must respond from their independent capacities for product development. Some competitors have ongoing development programs that provide them with product options to respond to the radical product innovation, and others may initiate such programs, but most commonly, competitors respond with variations or reformulations of their existing products (in our data, only 9.5% of all responding products contain new chemistry and merit priority review).

All these reported finding have implications for the timing of competitive response in the pharmaceutical industry (for a review, see Mathieu 2002). Competitors receive information about radical product innovations before the products are introduced, but the information is not reliable until the later stages of product development. When the information becomes reliable, competitors begin to respond, most commonly with reformulations of existing products. Reformulations have shorter development times than completely new products, but even so, there is some delay (see FDA 2005; Keyhani, Diener-West, and Powe 2006; Mathieu 2002). Overall, some competitive products may appear within months of the introduction of a radical product innovation, but as we indicated previously, the average response period is approximately one and a half years (Pharmaceutical Research and Manufacturers of America 2003), and some responses may even happen after the three-year observation window we use to measure response (see DiMasi and

FIGURE 3
Number of Radical Product Innovations per Category



Paquette 2004). In the latter case, the observations are right censored, and we account for this in our empirical model.[3]

Operationalizations and Data Sources for Hypothesized Variables

Introducer firm size. Sales, assets, and number of employees have all been used as proxies for firm size in prior strategy research, and these variables all tend to be highly correlated with one another (Agarwal 1979). We follow recent research in marketing strategy (Chandy and Tellis 1998) and use a firm's total sales in the year of product introduction to measure firm size.

It might be argued that the effects of the introducer's firm size are relative rather than absolute; that is, the likelihood of competitive response depends not so much on the absolute size of the introducing firm but rather on the relative sizes of the introducer and its competitors. The use of a relative measure is consistent with the effects of introducer firm size as discussed in prior literature, in which the presumed mechanism for the effect is the competitor's fear of retaliation (Robinson 1988; Shankar 1999). However, it is not consistent with our research, because our signaling arguments depend on the absolute (not relative) size of the introducing firm. Even so, the point is made moot by including competitor firm size in the analysis to capture any effects of the competitor's relative size.

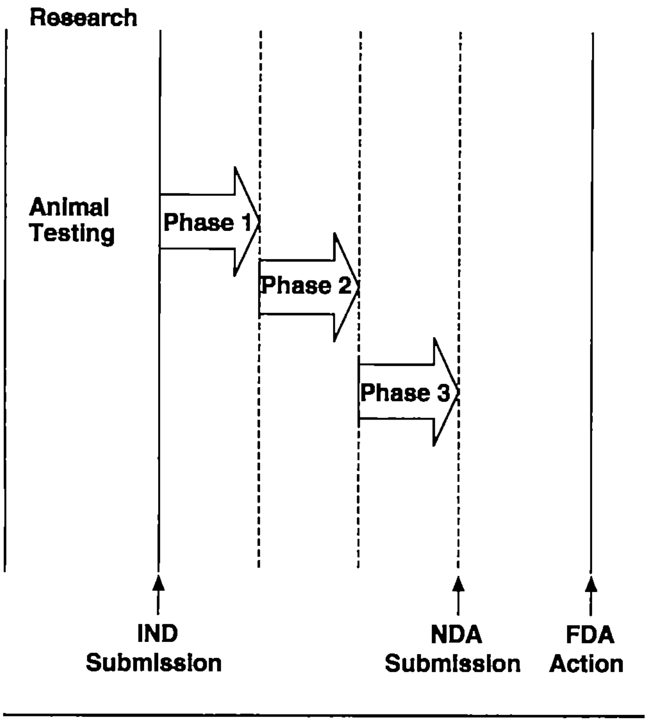
We obtained firm size data from four sources: (1) Standard & Poor's COMPUSTAT database, which includes financial information for companies that are publicly traded in the United States; (2) Thomson Datastream, which includes financial data for both U.S. and international companies; (3) Hoover's, which also covers some smaller and private companies; and (4) company Web sites.

Introducer market dependence. We used market dependence to signify the importance of a particular market to a firm. For the purpose of this study, we operationalized market dependence as the ratio of a firm's sales in a therapeutic product category to the firm's total sales in the year of introduction (for a similar measure in the airline industry, see Chen and Miller 1994). Virtually all the introducers derived at least some of their sales from the therapeutic categories in which they introduced the innovations; only 3 of the 52 radical product innovations were introduced by firms that previously derived no sales from the category in question. We obtained data on category sales for each firm from IMS

---

[3] Right censoring refers to a situation in which the firm might still be at risk of hazard (in this case, competitive response) at the conclusion of the observation period. It is incorrect to assume that such a firm will never respond, simply because it has not responded over the observation period.

FIGURE 4
New Drug Development Process



Source: Adapted from FDA (n.d.).
Notes: NDA = new drug application; IND = investigational new drug.

Health Global Services, which is a source of detailed market research data about the pharmaceutical industry.

Market size. We operationalized market size, which is hypothesized to interact with introducer firm size and introducer market dependence, as the total sales for the therapeutic category in which the radical product innovation was introduced, in the year of introduction (Reiffen and Ward 2005). We obtained market size data from IMS Health Global Services.

Operationalizations and Data Sources for Control Variables

Control variables we used in the analysis include competitor firm size, competitor market dependence, market growth, market size, market concentration, prior product introductions by the competitor, and order of response. We used these variables as controls because they could be related to competitive response for reasons discussed previously. We also used a control variable for whether the product is in the HIV therapeutic category. We used this variable because the FDA provided an automatic expedited review for HIV drugs during the measurement period, which may have affected the dynamics of innovation and response in this category. We already described the operationalization of market size; we operationalized the other control variables as follows:

- *Competitor firm size.* We measured firm size for each competitor in the same way and from the same data sources as introducer firm size.
- *Competitor market dependence.* We measured competitor market dependence in the same way and from the same data sources as introducer market dependence.
- *Market growth.* We operationalized market growth as the rate of change in market size in the year before introduction of the radical product innovation (relative to the previous year). We calculated growth rates from market size data obtained from IMS Health Global Services.
- *Market concentration.* We measured market concentration using the Herfindahl index, defined as the sum of squared market shares for the largest four competitors in a therapeutic category in the year the radical product innovation was introduced (Sorescu, Chandy, and Prabhu 2003). An alternative measure of market concentration—the sum of market shares of the top four firms in the market—provides similar results to those we report herein. We obtained market concentration data from IMS Health Global Services.
- *HIV category.* We included a dichotomous dummy variable that represents whether the radical product introduction belonged to the HIV category. We obtained relevant data from the NDA Pipeline.
- *Prior product introductions by the competitor.* We operationalized this variable as a dichotomous variable that represents whether a given competitor introduced a new product in the relevant product category during the three years before the introduction of a radical product innovation. We obtained data from the NDA Pipeline.
- *Order of response.* We measured order of response as the number of firms that already responded to the radical product innovation before response by a particular competitor. We obtained data from the NDA Pipeline.

Table 1 summarizes the descriptions of variables we used in this study, along with the data source for each variable. Table 2 provides descriptive statistics for the variables. Note that the variables are standardized before estimation.

Model Specification

We now turn to specification of the model used in our analysis. Our hypotheses address competitive response to radical product innovations; specifically, we focus on product introductions made by competitors in the three years following the introduction of a radical product innovation. This suggests three features that must be accounted for in any model specification. First, the response is discrete; a competitor either introduces a product in response or does not. Second, the response follows a temporal sequence; within our sample period, a firm could introduce at any time. Third, the data are right censored; a firm that did not introduce by the end of our sample period could still do so afterward. All these features suggest the appropriateness of using a hazard model specification to model our phenomenon of interest. Briefly, a hazard specification models the impact of a set of covariates on the probability of a discrete event (e.g., a product introduction) occurring.[4] Such models account for right censoring and the temporal nature of the

---

[4] Our interest in this article is on a single outcome—namely, whether firms respond with a product introduction. If there are multiple outcomes of interest, a competing risks modeling framework should be used (Lunn and McNeil 1995).

TABLE 1
Variables, Measures, and Sources

Variable	Operationalization	Source
Competitive response	Number of days until a focal competitor introduces a new product, starting from the day the radical innovation was introduced.	NDA Pipeline
Introducer size$_t$	Total dollar sales of the firm introducing the radical product innovation, in year t.	Standard & Poor's COMPUSTAT Thomson Datastream Hoover's Company Web sites
Introducer market dependence$_t$	Ratio of sales in a therapeutic category divided by the total sales for the introducing firm, in year t.	IMS Health Global Services
Market size$_t$	Total dollar sales of the therapeutic category in which the radical product innovation was introduced, in year t.	IMS Health Global Services
Competitor firm size$_t$	Total dollar sales of the focal competitor, in year t.	Standard & Poor's COMPUSTAT Thomson Datastream Hoover's Company Web sites
Competitor market dependence$_t$	Ratio of focal competitor's sales in a therapeutic category to the firm's total sales, in year t.	IMS Health Global Services
Market growth$_t$	Percentage change in the size of the market in year t – 1 compared with sales in year t – 2.	IMS Health Global Services
Market concentration$_t$	Sum of squared market shares for the largest four competitors in a therapeutic category in the year the radical product innovation was introduced.	IMS Health Global Services
HIV category	A dichotomous variable representing whether or not the radical product innovation was introduced in the HIV therapeutic category.	NDA Pipeline
Prior product introductions	A dichotomous variable representing whether a focal competitor introduced a new product in the three years preceding the introduction of a radical product innovation.	NDA Pipeline
Order of response	Number of firms that responded to the radical product innovation before response by the focal firm.	NDA Pipeline

TABLE 2
Descriptive Statistics

Variable	M	Minimum	Maximum
Competitive response (days)	790.68	.00	1096.00
Introducer size (in millions of dollars)	1,949.85	<.01	43,651.58
Introducer market dependence (%)	.20	<.01	1.00
Market size (in millions of dollars)	4,989.07	28.89	16,000.00
Competitor firm size (in millions of dollars)	2,041.74	<.01	43,651.58
Competitor market dependence (%)	.16	<.01	1.00
Market growth (%)	28.57	3.34	86.13
Market concentration (%)	.61	.30	1.00
Order of response (count)	4.38	.00	22.00

data naturally and have been widely used in marketing (e.g., Helsen and Schmittlein 1993; Jain and Vilcassim 1991; for an introductory primer, see Li 1995).

Having decided on a hazard specification, we needed to address several issues at the outset. Thus, consider a popular basic specification, the proportional hazard model

(Greene 2003; Seetharaman and Chintagunta 2003). The hazard function for a firm i can be written as follows:

$$(1) \qquad h(t|x_i) = h_0(t)\exp(\beta_0 + x_i\beta),$$

where $h(t|x_i)$ refers to the instantaneous hazard at time t, given a vector of covariates x_i; $h_0(t)$ refers to the baseline hazard rate; and β is a vector of unknown regression parameters. The hazard function "simply expresses the instantaneous probability of an event" (DuWors and Haines 1990, p. 487), or the "likelihood that an event that lasted until t will end in the next instance" (Grimshaw et al. 2005, p. 452). The choice of baseline hazard is crucial. A parametric specification could be used, and then either a specific distribution (e.g., exponential, Erlang 2) would be assumed or a flexible specification, such as the quadratic Box–Cox (Jain and Vilcassim 1991), would be chosen. Unfortunately, misspecified parametric specifications can lead to inconsistent estimates (Meyer 1995). The alternative, which we adopt, is to use a nonparametric specification (Cox 1972) and make no assumptions about the possible shape of the baseline hazard. (If the shape of the baseline hazard is of particular interest, parametric specifications can be compared with nonparametric specifications of the baseline hazard to determine which fits the data better; see, e.g., Seetharaman and Chintagunta 2003.) Because the intercept in the model in Equation 1 is unidentified from the baseline hazard in the Cox proportional hazard specification, we can rewrite the model as follows:

$$(2) \qquad h(t|x_i) = h_0(t)\exp(x_i\beta).$$

Although it is widely used and powerful, this specification does not account for unobserved heterogeneity (Gönül and Srinivasan 1993; Jain and Vilcassim 1991). It might well be that firms differ in ways that our covariates do not capture; for example, managerial ability is a variable that we do not include, but it could be argued that it might influence a firm's competitive response. It is well known that the exclusion of such factors could lead to a bias toward negative duration dependence (Heckman and Singer 1984). To account for this, we use a random-effects specification by introducing a new parameter, u, that varies across firms. In the hazard model literature (particularly in the biostastitics field; see, e.g., Vaupel, Manton, and Stallard 1979; and, more generally, Hougaard 2000), this is often referred to as a "frailty" model, and u is the frailty term. Formally, the specification now is as follows:

$$(3) \qquad h(t|x_i) = u_i h_0(t)\exp(x_i\beta).$$

Now, it is important to note that any firm in our data could have multiple incidences of product introductions (thus, the data are an unbalanced panel). Suppose that each firm i has J_i observations. With some abuse of notation,[5] we can write the hazard rate for the jth observation of the ith firm as follows (Fan and Li 2002):

$$h_{ij}(t|\mathbf{x}_{ij}, u_i) = h_0(t)u_i\exp(\mathbf{x}_{ij}'\beta);\ i = 1, \ldots, n;\ j = 1, \ldots, J_i,$$

where each firm i can be thought to constitute a "group" with j "members." The vector **x** contains all the independent variables discussed previously (i.e., introducer firm size, introducer market dependence, market size, and other control variables).

We still need to specify the distribution of u. Following common practice in the literature (Clayton 1978; Hougaard 2000; Vaupel, Manton, and Stallard 1979), we specify a gamma distribution[6] for this random variable, with a mean of 1 and a variance of $1/\theta$; that is,

$$(4) \qquad g(u) = \frac{u^{\theta - 1}\theta^{\theta}\exp(-u\theta)}{\Gamma(\theta)}.$$

The next step is to obtain the likelihood function to be maximized. It is a standard result that the likelihood function is the product of the density function for uncensored observations and the survivor function for censored observations. Applied here, we obtain the following:

$$(5) \qquad \prod_{i=1}^{n}\prod_{j=1}^{J_i}\left\{\left[h(z_{ij}|\mathbf{x}_{ij}, u_i)\right]^{\delta_{ij}} S(z_{ij}|\mathbf{x}_{ij}, u_i)\right\}\prod_{i=1}^{n} g(u_i),$$

where $S(\cdot)$ represents the conditional survivor function, δ is a censoring indicator (i.e., 1 if censored), and z is the observed time (more precisely, $z = \min\{t, c\}$, where t and c are the survival and censoring times, respectively). Integrating the heterogeneity u out of Equation 5 gives us the likelihood of the observed data, which is maximized to obtain the parameters θ (capturing frailty) and the vector β:

$$L(\beta, \theta, H) = \exp\left[\beta'\left(\sum_{i=1}^{n}\sum_{j=1}^{J_i}\delta_{ij}\mathbf{x}_{ij}\right)\right]$$

$$\times\prod_{i=1}^{n}\frac{\theta^{\theta}\prod_{j=1}^{J_i}\left[h_0(z_{ij})\right]^{\delta_{ij}}}{\Gamma(\theta)\left[\sum_{j=1}^{J_i}H_0(z_{ij})\exp(\mathbf{x}_{ij}'\beta) + \theta\right]^{(A_i + \theta)}},$$

where $A_i = \sum_{j=1}^{J_i}\delta_{ij}$.

Results

A Wald test for the overall fit of the model yields a chi-square value of 363.13, which is significant at the $p < .01$ level. Table 3 reports the estimated coefficients. A comparison of the full model (with main effects and interactions)

---

[5] In general, this notation would be used in the case in which there were j individuals as part of a group i, sharing a common frailty u_i. In such a case, the frailty would be a random term multiplicatively affecting the hazard rate of all members of the group.

[6] Use of an inverse Gaussian distribution for the frailty term (Hougaard 2000) yields similar results in our empirical context.

TABLE 3
Hazard Model Results

	Main Effects–Only Model		Main + Interaction Effects Model	
Hypothesized Variables	**Coefficient**	**SE**	**Coefficient**	**SE**
Introducer firm size	.55*	.08	.64*	.11
Introducer market dependence	.34*	.05	.52*	.07
Introducer firm size × market size	—	—	–.55*	.16
Introducer market dependence × market size	—	—	.03	.09
Control Variables	**Coefficient**	**SE**	**Coefficient**	**SE**
Competitor firm size	1.18*	.14	1.15*	.14
Competitor market dependence	.43*	.05	.44*	.05
Market size	.09	.06	.24*	.06
Market growth	.01*	2.69×10^{-3}	.01*	2.69×10^{-3}
Market concentration	.10	.06	.70	.49
HIV category	.08	.16	.06	.16
Prior product introductions	–.19*	.06	–.18*	.06
Order of response	–1.04*	.10	–1.13*	.10
Frailty parameter	.23*	.08	.19*	.073
Log-likelihood value	–1908.58		–1896.49	
Wald χ^2 statistic	340.11*		348.91*	
AIC	3839.16		3818.98	
BIC	–1924.28		–1915.04	

*$p < .01$.

Notes: The Akaike information criterion (AIC) is given as –2L + 2k, where L is the log-likelihood function and k is the number of parameters. Models with lower values of AIC are preferred. The Bayesian information criterion (BIC) is given as L – .5 × k × ln(n), where n is the number of observatons. Models with lower values of BIC are preferred.

with a model with main effects only indicates that it is appropriate to include the interaction effects in the model. Coefficients for almost all predictor variables are significantly different from zero.

Results of Hypothesis Tests

H_1 suggested that the larger the firm introducing a radical product innovation, the greater is the likelihood of competitive response. The effect is significant and in the expected direction ($\beta = .64$, $p < .01$), in support of H_1. Although the magnitude of the coefficient is not immediately interpretable, we can convert it to a hazard ratio, which is the exponential of the coefficient. A hazard ratio of 1 suggests that the variable has no impact on the probability of responding; a ratio greater than 1 suggests that it positively influences the probability of response, and a ratio less than 1 suggests a negative influence. The more the ratio departs from 1, the stronger is the effect. The hazard ratio for introducer firm size is 1.91, which implies that at any given time, a one-standard-deviation increase in the size of the introducing firm almost doubles the chance that any given competitor will respond in the next period if it has not responded already.

This finding reverses prior results obtained in the context of incremental innovations, in which introducer size has been found to be negatively related to competitive response (Bowman and Gatignon 1995; Shankar 1999). This reversal of effects lends credibility to the idea that competitive dynamics for radical product innovations may qualitatively differ from those for incremental innovations.

H_2 argued that the more market dependent the firm introducing a radical product innovation, the greater is the likelihood of competitive response. The effect is significant and in the hypothesized direction ($\beta = .52$, $p < .01$), in support of H_2. The hazard ratio is 1.69, which suggests that increasing market dependence by one standard deviation increases the probability of responding 1.69 times.

H_{3a} predicted that market size would negatively moderate the influence of introducer size on the likelihood of competitive response, such that as the market becomes larger, the effect of introducer size becomes smaller (and as the market becomes smaller, the effect of introducer size becomes larger). The hypothesis is supported ($\beta = -.55$, $p < .01$). The hazard ratio is .57, which suggests that an increase in market size by one standard deviation cuts the effect of introducer size almost by half.

H_{3b} predicted that market size would negatively moderate the influence of introducer market dependence on the likelihood of competitive response, such that as the market becomes smaller, the effect of introducer market dependence becomes larger. The observed effect is not significant ($\beta = .03$, $p > .1$), with a hazard ratio of 1.03.

Overall, three of the four hypotheses are confirmed. In general, the results are consistent with our argument that competitive response to radical product innovations is higher when some aspect of the product introduction, such as the introducer firm size, introducer market dependence, or the interaction of market size with introducer firm size, provides competitors with signals that the innovation is likely to increase the size of the market.

Other Results

Competitor firm size. Competitor firm size has a significant, positive impact on the likelihood of response ($\beta = 1.15$, $p < .01$). This is a strong effect; the hazard ratio of 3.17 suggests that a one standard deviation increase in the competitor's size produces an almost threefold increase in the probability of reacting at any given time. This result contradicts prior findings in the context of incremental innovations that show either a negative impact (Kuester, Homburg, and Robertson 1999; Shankar 1999) or no impact (Robinson 1988) of competitor firm size on response. The general explanation for prior findings has been that larger competitors are less likely to respond because of bureaucratic inflexibility or inertia (Tornatzky and Fleischer 1990). However, it can also be argued that larger firms have greater resources and, thus, greater capabilities for response (Chandy, Prabhu, and Antia 2003; Chandy and Tellis 1998; Sorescu, Chandy, and Prabhu 2003), and in the context of radical product innovations, in which the destabilizing potential of the innovation may provide an inherent motivation to respond, these capabilities may result in larger firms being more likely to respond. Our findings are consistent with this latter argument, which again suggests that there could be qualitative differences in competitive dynamics for radical and incremental innovations.

Competitor market dependence. Competitor market dependence also has a significant, positive impact on response ($\beta = .44$, $p < .01$). The hazard ratio is 1.55. This finding is what would be expected; namely, market-dependent competitors have a greater stake in the market, which motivates a stronger defensive posture and a greater likelihood of response.

Market size. A larger category elicits a greater likelihood of response ($\beta = .24$, $p < .01$). The hazard ratio is 1.27. This finding supports the idea that competitors respond more aggressively in larger markets.

Market growth rate. A faster-growing category elicits a greater likelihood of response ($\beta = .01$, $p < .01$). The effect is weak, as is shown by a hazard ratio of 1.01. Prior empirical results for this variable have been mixed (Shankar 1999). This result, combined with the positive result for market size, supports prior studies that suggest that competitive response is stronger in growing markets because competitors view the market as more attractive and are willing to fight for it (Bowman and Gatignon 1995; Kuester, Homburg, and Robertson 1999).

Market concentration. Market concentration has no significant impact on the likelihood of response ($\beta = .70$, $p > .1$). The hazard ratio is 1.08.

HIV category. A product introduction in the HIV category has no significant affect on the likelihood of response ($\beta = .06$, $p > .1$). The hazard ratio is 1.06.

Prior product introductions. Competitors that introduced a new product in the same product category within the previous three years are significantly less likely to respond to a radical product introduction ($\beta = -.18$, $p < .01$). The hazard ratio is .32.

Order of response. The greater the number of competitors that have already responded to a radical product introduction, the less likely it is that a given firm will respond with a product introduction ($\beta = -1.13$, $p < .01$). The hazard ratio is .83.

Robustness Checks

We conducted several robustness checks. First, we examined whether our results were sensitive to the cutoff date in our sample (i.e., whether changing the extent of right censoring affects the results). To do this, we estimated a Cox hazard model with the cutoff date in our sample changed from three years to two years. Table 4, Panel A, reports the results of this change in the degree of right censoring. The results are qualitatively unchanged. In the same vein, we estimated a discrete choice probit model that does not account for right censoring at all (and treats the dependent variable purely as occurrence of response without regard to time). Table 4, Panel B, reports the results of this analysis; the results are qualitatively unchanged.

Second, we checked the robustness of our specification by trying alternative parametric specifications of the hazard—specifically, a Weibull hazard with gamma frailty and a log-logistic model with gamma frailty. The results are similar in both cases and are similar to the Cox proportional hazard model we reported.

Third, we examined robustness to the particular sample used by conducting a bootstrapping analysis with 50 repetitions (Table 5). Again, the results are largely unchanged.

Fourth, we checked the predictive validity of our specification (Srinivasan, Lilien, and Rangaswamy 2006). To do this, we randomly picked two-thirds of our sample as the estimation sample and the remaining one-third as the holdout sample. Using estimates from the estimation sample, we calculated the hazard ratio for each observation in the holdout sample. If the hazard ratio exceeded the baseline rate, we counted that observation as having experienced the hazard (i.e., responded to the product introduction). We then compared this with the actual response to determine whether the model predicted correctly, and we computed the hit rate over all observations. The hit rate on the holdout sample is 83.35%; this high figure provides reassurance about the predictive validity of our model. Delving further, we find that of a total holdout sample of 245, the model predicts a response in 149 cases, 113 of which indeed responded, and no response in 96 cases, 94 of which indeed did not respond. This implies a false-positive rate of .22 and a false-negative rate of .02. Using the approach that Morrison (1969) suggests for evaluating the predictive validity of a discriminant classifying function, we find that our .83 hit rate compares favorably with .49, the hit rate that would have been obtained by chance: $c_{pro} = \alpha p + (1 - \alpha)(1 - p)$, where p = true proportion of responders and α = proportion of responders predicted by the model.

Discussion

In this article, we build a middle-range theory (Bourgeois 1979) of competitive responses to radical product innovation. We introduce the concepts of market expansion and entry thresholds to develop and test new hypotheses about competitive response to radical product innovation. Our analyses reveal some novel findings. Innovations introduced by large firms and by market-dependent firms are especially likely to elicit a greater competitive response. The likelihood of competitive response is the greatest when large firms enter small markets. The subsequent paragraphs discuss the implication of this research for research and practice.

TABLE 4
Robustness Checks

A: Robustness to Degree of Right Censoring

Hypothesized Variables	Estimate	SE
Introducer firm size	.18**	.05
Introducer market dependence	.26**	.05
Introducer firm size × market size	–.21*	.09
Introducer market dependence × market size	.03	.06
Control Variables	**Estimate**	**SE**
Competitor firm size	.26**	.05
Competitor market dependence	.14**	.04
Market size	.11*	.04
Market growth	.01**	.002
Market concentration	.03	.04
HIV category	–.04	.13
Prior product introductions	–.25**	.08
Order of response	–.10**	.01
Frailty parameter	.04**	.02
Log-likelihood value	–3650.81	
Wald χ^2 statistic	210.90**	

B: Probit Model Estimates (Dependent Variable: Response)

Variable	Estimate	SE
Introducer firm size	.486**	.121
Introducer market dependence	.492**	.120
Introducer firm size × market size	–.744**	.229
Introducer market dependence × market size	.182	.143
Competitor firm size	1.635**	.187
Competitor market dependence	1.137**	.124
Market size	.210*	.104
Market growth	.017**	.004
Market concentration	.035	.086
Prior product introductions	–.418**	.093
Order of response	–.588**	.115
HIV category	.106	.266
Intercept	–.956**	.227

*$p < .05$.
**$p < .01$

Know Thyself to Predict Thy Opponent

Modern marketing science provides managers with sophisticated techniques to predict the likely outcomes of product introduction. The richness of the consumer response models used in these projections contrasts sharply with the sketchiness of the competitor response models used. Whereas consumer response models are often calibrated with detailed empirical analyses of responses from current and future customers, competitor response models are often ad hoc and based on subjective judgments (Hauser, Tellis, and Griffin 2006).

The results from this research suggest a need for greater care and more careful anticipation when calibrating the likelihood of competitor response. For firms introducing radical product innovations, these results highlight the usefulness of looking in the mirror for clues regarding the likely responses of their competitors. By examining the signals they send about their own entry thresholds, firms can assess the likelihood and speed of competitive response to their radical product innovations. Large size and market dependence tend to signal high entry thresholds; as such, large and market-dependent firms should anticipate and account for greater competitive response to their radical product innovations.

Watch for Trickles That Can Turn into Gushers

Most analyses of the likelihood of competitive response to a particular market take the size of the market as a given. Therefore, many current analyses imply that competitive response is likely to be limited when the market being entered is small (see Montaguti, Kuester, and Robertson 2002). The results from this article suggest a need to rethink this conclusion.

TABLE 5
Bootstrap Analysis

Hypothesized Variables	Estimate	SE
Introducer firm size	.65**	.11
Introducer market dependence	.53**	.076
Introducer firm size × market size	–.56**	.21
Introducer market dependence × market size	.01	.12
Control Variables	**Estimate**	**SE**
Competitor firm size	1.15**	.13
Competitor market dependence	.43**	.08
Market size	.25**	.10
Market growth	.01**	.004
Market concentration	.08	.06
HIV category	.06	.14
Prior product introductions	–.37*	.16
Order of response	–.20**	.03
Frailty parameter	.04**	.02
Log-likelihood value	–3650.81	
Wald χ^2 statistic	338.98**	

*$p < .05$.
**$p < 01$

Our results emphasize the need to examine markets dynamically, not just by considering them as they are today but as they could be tomorrow. Radical product innovations can cause previously small markets to explode in size. Visions of explosive growth inevitably attract competitive response, even before such growth has actually taken place. Our results indicate that competitors use the entry threshold of the firm that introduces the radical product innovation in a market as a signal of the eventual size of the market.

Do Not Assume That Entry by Giants Will Dissuade Others from Entering

Conventional wisdom suggests that entry into a market by large firms will lead to lower entry by competitors because of fears of retaliation (Robinson 1988; Shankar 1999). This logic would lead to the conclusion that large firms can be sanguine about competition in such cases.

Our results suggest that in the context of radical product innovation, this conclusion should be just the opposite. Specifically, our results suggest that the introduction of a radical product innovation by a large firm is likely to lead to a surge, not a reduction, in competitive response. Visions of market growth that are triggered by the introduction of a radical product innovation by a large firm may swamp any threats of retaliatory behavior. The surge in competitive response is especially prominent in cases in which the large firm enters a small market because the signal regarding market growth is especially powerful in such cases. Our results caution managers of large and market-dependent firms against being overly optimistic about sustaining monopoly profits from radical product innovation.

Limitations and Further Research

This research has several limitations, and some of these can provide avenues for further research. Although testing our model within the context of a single industry and using the U.S. pharmaceutical industry in particular have advantages, we acknowledge the specific characteristics of this industry, including heavy regulation, preannouncement of innovation research, the particular speed of innovation cycles, and a mix of firm sizes ranging from extremely large multinational corporations to small companies. We believe that confidence can be placed in the direction of the effects we observed in this research because most of the effects are predicted by our hypotheses or replicate prior research (regarding the role of theory and replication in research generalization, see Blair and Zinkhan 2006). However, industry-specific effects might limit the generalizability of effect sizes, so it would be useful to extend this line of research to include other industries that vary in regulatory processes (e.g., international pharmaceutical markets, nonprescription drug markets, various unregulated markets), innovation cycles (e.g., semiconductors, electronics, paint), or concentration (e.g., aircraft manufacture).

In addition, we consider only one dimension of response, product response. Response can also occur on other dimensions of the marketing-mix variables. For example, firms can respond with advertising increases or price cuts. Although existing research (e.g., Kuester, Homburg, and Robertson 1999) suggests that responses to actions on the product dimension are highly likely to be on the product dimension as well (and especially so for radical product innovations), it would nevertheless be worthwhile to examine, for example, how large versus small firms behave on other marketing-mix dimensions when reacting to the introduction of a radical product innovation.

It would also be a worthwhile extension to examine how well products introduced as responses to radical innovations fare in the market. Are they more or less likely to succeed than other product introductions? Finally, as is common with most studies that use secondary data, we inferred firm decision making from observed outcomes. It would be useful to supplement our study with primary data obtained from managers to examine how decision making is influenced by the variables we suggested and to examine whether there are any important variables our study omitted.

REFERENCES

Agarwal, N.C. (1979), "On the Interchangeability of Size Measures," *Academy of Management Journal*, 22 (2), 404–409.

Axelrod, Robert (2002), *The Evolution of Cooperation*. New York: HarperCollins.

Bernheim B. Douglas and Michael D. Whinston (1998), "Incomplete Contracts and Strategic Ambiguity," *American Economic Review*, 88 (4), 902–933.

Blair, Ed and George Zinkhan (2006), "Nonresponse and Generalizability in Academic Research," *Journal of the Academy of Marketing Science*, 34 (1), 4–7.

Blundell, Richard, Rachel Griffith, and John Van Reenen (1999), "Market Share, Market Value, and Innovation in a Panel of British Manufacturing Firms," *Review of Economic Studies*, 66 (228), 529–54.

Bourgeois, L.J. (1979), "Toward a Method of Middle-Range Theorizing," *Academy of Management Review*, 4 (3), 443–47.

Bowman, Douglas and Hubert Gatignon (1995), "Determinants of Competitor Response Time to a New Product Introduction," *Journal of Marketing Research*, 32 (February), 42–53.

Bresnahan, T.F. and P.C Reiss (1991), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99 (5), 977–1009.

Chandy, Rajesh K., Jaideep C. Prabhu, and Kersi D. Antia (2003), "What Will the Future Bring? Technology Expectations, Dominance, and Radical Product Innovation," *Journal of Marketing*, 66 (July), 1–18.

——— and Gerard J. Tellis (1998), "Organizing for Radical Product Innovation: The Overlooked Role of Willingness to Cannibalize," *Journal of Marketing Research*, 35 (November), 474–87.

——— and ——— (2000), "The Incumbent's Curse? Incumbency, Size, and Radical Product Innovation," *Journal of Marketing*, 64 (July), 1–17.

Chen, Ming-Jer and Ian C MacMillan (1992), "Nonresponse and Delayed Response to Competitive Moves: The Roles of Competitor Dependence and Action Irreversibility," *Academy of Management Journal*, 35 (3), 539–70.

——— and Danny Miller (1994), "Competitive Attack, Retaliation and Performance: An Exploratory-Valence Framework," *Strategic Management Journal*, 15 (February), 85–102.

Christensen, Clayton M. (1997), *The Innovator's Dilemma*. New York: Harper Business.

Clayton, D.G. (1978), "A Model for Association in Bivariate Life Tables and Its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence," *Biometrika*, 65 (1), 141–51.

Cox, David R. (1972), "Regression Models and Life-Tables," *Journal of the Royal Statistical Society: Series B*, 34 (2), 187–200.

DiMasi, Joseph A. and Cherie Paquette (2004), "The Economics of Follow-On Drug Research and Development," *Pharmacoeconomics*, 22 (2), 1–14.

DuWors, Richard E., Jr., and George H. Haines Jr. (1990), "Even History Analysis Measures of Brand Loyalty," *Journal of Marketing Research*, 27 (November), 485–93.

Fan, Jianqing and Runze Li (2002), "Variable Selection for Cox's Proportional Hazards Model and Frailty Model," *Annals of Statistics*, 30 (1), 74–99.

FDA (1999), *From Test Tube to Patient: Improving Health Through Human Drugs*. Washington, DC: U.S. Government Printing Office.

——— (2005), *FY 2005 Performance Report to the President and the Congress for the Prescription Drug User Fee Act*. Washington, DC: U.S. Government Printing Office.

——— (n.d.), "The New Drug Development Process· Steps from Test Tube to New Drug Application Review," (accessed January 7, 2008), [available at http://www.fda.gov/cder/handbook/develop.htm].

Golder, Peter N. and Gerard J. Tellis (1997), "Will It Ever Fly? Modeling the Takeoff of New Consumer Durables," *Marketing Science*, 16 (3), 256–70.

Gönül, Füsun and Kannan Srinivasan (1993), "Consumer Purchase Behavior in a Frequently Bought Product Category: Estimation Issues and Managerial Insights from a Hazard Function Model with Heterogeneity," *Journal of the American Statistical Association*, 88 (424), 1219–27.

Govindarajan, Vijay and Praveen K. Kopalle (2004), "How Incumbents Can Introduce Radical and Disruptive Innovations: Theoretical and Empirical Analyses," working paper, Tuck School of Business, Dartmouth College.

Greene, William (2003), *Econometric Analysis*. Englewood Cliffs, NJ: Prentice Hall.

Grimshaw, Scott D., James McDonald, Grant R. McQueen, and Steven Thorley (2005), "Estimating Hazard Functions for Discrete Lifetimes," *Communications in Statistics-Simulation and Computation*, 34 (2), 451–63.

Gruca, Thomas S., K. Ravi Kumar, and D. Sudharshan (1992), "An Equilibrium-Analysis of Defensive, Response to Entry Using a Coupled Response Function Model," *Marketing Science*, 11 (4), 348–58.

——— and D. Sudharshan (1995), "A Framework for Entry Deterrence Strategy: The Competitive Environment, Choices, and Consequences," *Journal of Marketing*, 59 (July), 44–55.

Hauser, John, Gerard Tellis, and Abbie Griffin (2006), "Research on Innovation: A Review and Agenda for *Marketing Science*," *Marketing Science*, 25 (6), 687–717.

Heckman, James and Burton Singer (1984), "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models of Duration Data," *Econometrica*, 51 (2), 271–320.

Heide, Jan B. and Allen M. Weiss (1995), "Vendor Consideration and Switching Behavior in High-Technology Markets," *Journal of Marketing*, 59 (July), 30–43.

Heil, Oliver and Arien W. Langvardt (1994), "The Interface Between Competitive Market Signaling and Antitrust Law," *Journal of Marketing*, 58 (July), 81–96.

——— and Thomas S. Robertson (1991), "Toward a Theory of Competitive Market Signaling: A Research Agenda," *Strategic Management Journal*, 12 (6), 403–18.

Helsen, Kristiaan and David Schmittlein (1993), "Analyzing Duration Times in Marketing: Evidence for the Effectiveness of Hazard Rate Models," *Marketing Science*, 12 (4), 395–414.

Hougaard, P. (2000), *Analysis of Multivariate Survival Data*. New York: Springer.

IMS Health Global Services (2004), "IMS National Sales Perspectives," research report.

Jain, Deepak and Naufel Vilcassim (1991), "Modeling Purchase-Timing and Brand-Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity," *Journal of Marketing Research*, 28 (February), 29–41.

Kerin, Roger A., Michael G. Harvey, and James T. Rothe (1978), "Cannibalism and New Product Development," *Business Horizons*, (October), 25–31.

Keyhani, Salomeh, Marie Diener-West, and Neil Powe (2006), "Are Development Times for Pharmaceuticals Increasing or Decreasing?" *Health Affairs*, 25 (2), 461–68.

Kuester, Sabine, Christian Homburg, and Thomas S. Robertson (1999), "Retaliatory Behavior to New Product Entry," *Journal of Marketing*, 63 (October), 90–106.

Li, Shaomin (1995), "Survival Analysis," *Marketing Research*, 7 (4), 17–23.

Lunn, Mary and Don McNeil (1995), "Applying Cox Regression to Competing Risks," *Biometrics*, 51 (2), 524–32.

Mahajan, Vijay, Subhash Sharma, and Robert D. Buzzell (1993), "Assessing the Impact of Competitive Entry on Market Expansion and Incumbent Sales," *Journal of Marketing*, 57 (July), 39–52.

Mathieu, Mark P. (2002), *New Drug Development: A Regulatory Overview*. Waltham, MA: Parexel International Corporation.

Meyer, Bruce (1995), "Semiparametric Estimation of Hazard Models," working paper, Department of Economics, Northwestern University, and National Bureau Economic Research.

Min, Sungwook, Manohar U. Kalwani, and William T. Robinson (2006), "Market Pioneer and Early Follower Survival Risks: A Contingency Analysis of Really New Versus Incrementally New Product-Markets," *Journal of Marketing Research*, 70 (February), 15–33.

Montaguti, Elisa, Sabine Kuester, and Thomas Robertson (2002), "Entry Strategy for Radical Product Innovations: A Conceptual Model and Propositional Inventory," *International Journal of Research in Marketing*, 19 (1), 21–42.

Moorthy, Sridhar (1985), "Using Game Theory to Model Competition," *Journal of Marketing Research*, 22 (August), 262–82.

Morrison, Donald (1969), "On the Interpretation of Discriminant Analysis," *Journal of Marketing Research*, 6 (May), 153–63.

Nickell, Steven (1996), "Competition and Corporate Performance," *Journal of Political Economy*, 54 (4), 724–46.

Nijssen, Edwin, Bas Hillebrand, and P.A.M. Vermeulen (2005), "Unraveling Willingness to Cannibalize: A Closer Look at the Barrier to Radical Innovation," *Technovation*, 25 (12), 1400–1409.

Pharmaceutical Research and Manufacturers of America (2003), *Pharmaceutical Industry Profile 2003*. Washington, DC: PhRMA.

Porter, Michael E. (1980), *Competitive Strategy*. New York: The Free Press.

Prabhu, Jaideep and David W. Stewart (2001), "Signaling Strategies in Competitive Interaction: Building Reputations and Hiding the Truth," *Journal of Marketing Research*, 38 (February), 62–72.

Reiffen, David and Michael R. Ward (2005), "Generic Drug Industry Dynamics," *Review of Economics and Statistics*, 87 (1), 37–49.

Rhoades, Stephen (1973), "The Impact of Diversification on Industry Profit Performance in 241 Manufacturing Industries: 1963," *Review of Economics and Statistics*, 55 (2), 146–55.

Robertson, Thomas, Jehoshua Eliashberg, and Talia Rymon (1995), "New Product Announcement Signals and Incumbent Reactions," *Journal of Marketing*, 59 (July), 1–15.

Robinson, William T. (1988), "Marketing Mix Reactions to Entry," *Marketing Science*, 7 (Fall), 368–92.

Rust, Roland T. (2006), "The Maturation of Marketing as an Academic Discipline," *Journal of Marketing*, 70 (August), 1–2.

Scherer, F.M (2000), "The Pharmaceutical Industry and World Intellectual Property Standards," *Vanderbilt Law Review*, 53 (6), 2245–54.

Schumpeter, J.A. (1942), *Capitalism, Socialism, and Democracy*. New York: Harper.

Seetharaman, P.B. and Pradeep K. Chintagunta (2003), "The Proportional Hazard Model for Purchase Timing: A Comparison of Alternative Specifications," *Journal of Business and Economic Statistics*, 21 (3), 368–83.

Shankar, Venkatesh (1997), "Pioneers' Marketing Mix Reaction to Entry in Different Competitive Game Structures: Theoretical Analysis and Empirical Illustration," *Marketing Science*, 16 (3), 271–93.

——— (1999), "New Product Introduction and Incumbent Response Strategies: Their Interrelationship and the Role of Multimarket Contact," *Journal of Marketing Research*, 36 (August), 327–44.

Sheremata, Willow (2004), "Competing Through Innovation in Network Markets: Strategies for Challengers," *Academy of Management Review*, 29 (3), 359–77.

Sorescu, Alina, Rajesh Chandy, and Jaideep Prabhu (2003), "Sources and Financial Consequences of Radical Innovation: Insights from Pharmaceuticals," *Journal of Marketing*, 66 (October), 82–102.

Spence, A. Michael (1974), *Market Signaling*. Cambridge, MA: Harvard University Press.

Srinivasan, Raji, Gary Lilien, and Arvind Rangaswamy (2006), "The Emergence of Dominant Designs," *Journal of Marketing*, 70 (April), 1–17.

Tellis, Gerard J. and Peter N. Golder (2001), *Will and Vision: How Latecomers Grow to Dominate Markets*. New York: McGraw-Hill.

Tornatzky, Louis G. and Mitchell Fleischer (1990), *The Process of Technological Innovation*. Lexington, MA: Lexington Books.

Van Heerde, Harald, Carl Mela, and Puneet Manchanda (2004), "The Dynamic Effect of Innovation on Market Structure," *Journal of Marketing Research*, 41 (May), 166–83.

Vaupel, J.W., K.G. Manton, and E. Stallard (1979), "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality," *Demography*, 16 (3), 439–54.

Weitz, Bart (1985), "Introduction to the Special Issue on Competition in Marketing," *Journal of Marketing Research*, 22 (August), 229–36.

Wuyts, Stefan, Stefan Stremersch, and Shantanu Dutta (2004), "Portfolios of Interfirm Agreements in Technology-Intensive Markets: Consequences for Innovation and Profitability," *Journal of Marketing*, 68 (April), 88–100.

Yeoh, Poh-Lin and Kendall Roth (1999), "An Empirical Analysis of Sustained Advantage in the U.S. Pharmaceutical Industry: Impact of Firm Resources and Capabilities," *Strategic Management Journal*, 20 (7), 637–53.

Yip, G.S. (1982), *Barriers to Entry: A Corporate-Strategy Perspective*. Lexington, MA: D.C. Heath and Company.

Jing Lei, Niraj Dawar, & Jos Lemmink

Negative Spillover in Brand Portfolios: Exploring the Antecedents of Asymmetric Effects

Marketers cultivate brand relatedness in their brand portfolios to increase marketing efficiency through positive spillover of brand equity. However, creating linkages between brands may also make them vulnerable to negative spillover. This research investigates the structure of relatedness in a brand portfolio to understand the nature of spillover effects. The results of two experiments show that the magnitude of spillover between brands is a function of not only the strength of brand associations but also their directionality. The results also show that the directional strength of association is influenced by the number and salience of associations linked to each brand. The authors draw implications for a theoretical understanding of spillover effects in marketing, as well as for the management of brand portfolios.

Keywords: product-harm crises, spillover effects, association asymmetry, brand portfolio

The conventional single-brand company and standalone brands are increasingly rare (Laforet and Saunders 1999). Complex brand portfolios with multiple brands, subbrands, endorsed brands, and cobrands are ubiquitous because brand relatedness helps sellers fulfill market segmentation, brand strategy, and marketing efficiency objectives (Aaker and Joachimsthaler 2000). Relatedness is cultivated through a variety of marketing means, including the use of common brand names and logos, similar trade dress and design, similar or related advertising, promotion of complementary usage, and even proximate shelf location (Aaker and Joachimsthaler 2000; Rajagopal and Sanchez 2004). After they are established, consumers' perceptions of brand relatedness influence the processing of brand-related external information. Indeed, a brand cannot be isolated from exposure to information about related brands; inevitably, it becomes subject to indirect effects. For the most part, these indirect effects are intended to be positive, as when affect associated with a well-known brand spills over onto brands within the same brand portfolio. However, increased brand relatedness may also expose brands to unintended risk from related brands in a portfolio when negative incidents (e.g., product-harm crises) occur. For researchers and marketers alike, the implication is that neglecting indirect effects on brands limits the understanding of the antecedents of important marketing constructs, such as brand evaluations and purchase likelihood.

Spillover from extension products to parent brand evaluations (e.g., John, Loken, and Joiner 1998) and from one brand to competing brands within the same product category (Roehm and Tybout 2006) has been previously documented. Such studies demonstrate that the strength of linkages between brands and their products or product category is a good predictor of the magnitude of spillover. However, brand linkages can be characterized by more than just the strength of association. Associative network theory (e.g., Collins and Loftus 1975) suggests that linkages between two concepts can point in both directions. Taking the directionality of linkages into account suggests a pattern of spillover that cannot be predicted on the basis of nondirectional strength of association alone. For example, negative information about Special K breakfast cereal might not affect evaluations of Corn Flakes to the same extent as the same information about Corn Flakes affects evaluations of Special K. In particular, it becomes necessary to spell out which of the two directional linkages between brands predicts the spillover from one to the other. From a marketing standpoint, asymmetric strength of association suggests asymmetric vulnerability of brands in a portfolio. Ignoring the directional nature of brand associations could lead to inaccurate predictions and expose brands to unintended risk if the incorrect direction of the linkage is used to predict spillover. However, the directionality of brand associations has been only conceptually proposed (e.g., Farquhar and Herr 1993) but never empirically examined, nor has the impact of association asymmetry on spillover effects received much research attention.

Jing Lei is an assistant professor, Faculty of Business and Information Technology, University of Ontario Institute of Technology (e-mail: jill.lei@uoit.ca). Niraj Dawar is R.A. Barford Professor of Marketing, Ivey Business School, University of Western Ontario (e-mail: ndawar@ivey.uwo.ca). Jos Lemmink is Professor of Marketing, Department of Marketing, Universiteit Maastricht (e-mail: j.lemmink@mw.unimaas.nl). The authors thank Pierre Chandon, June Cotte, Robert Fisher, Shanker Krishnan, Donald Lehmann, Kyle Murray, and the three anonymous *JM* reviewers for their valuable comments on previous drafts of this article. The authors also thank Leon van Poelwijk and Marjolein Bóggemann for their help in the data collection process. This research was supported in part by a grant to the second author from the Social Sciences and Humanities Research Council of Canada.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

If the directionality of association indeed influences the pattern of spillover, the antecedents of asymmetry deserve attention because these antecedents determine the network structure that guides the subsequent processing of external information and its indirect effects on brands. Previous research has primarily examined brand-related linkages as part of category structures (e.g., Farquhar, Herr, and Fazio 1990). It suggests that brand–category linkages can be asymmetric in strength and that this asymmetry influences how consumers process brand-related information. However, little is known about the drivers of association asymmetry or whether asymmetry indeed exists in linkages between brands in a brand portfolio. An understanding of the structure of interbrand relatedness would not only sharpen our predictions about spillover but also provide guidance in the design of brand portfolio architecture.

The objectives of this article are to examine the impact of both the strength and the directionality of association on spillover, as well as the antecedents of association asymmetry. A typical brand portfolio consists of a parent brand and subbrands as nodes and linkages of varying strength between these entities as the basic elements. We examine spillover by introducing negatively valenced information (i.e., product-harm crises) that implicates different brand entities within the portfolio and by predicting and observing the indirect negative consequences on related brands. Findings from two experiments show that spillover is influenced by directional strength of association between brands and that the directional strength of association is a function of the number and salience of associations linked to each brand. From these results, we draw theoretical implications for spillover effects in marketing and suggestions for the management of brand portfolios.

Cognitive Brand Networks

Our examination of the impact of external information on consumers' brand evaluations begins with the premise that brand representations are not cognitively independent. However, this premise still leaves open the choice of representational theory and the nature of brand associations that we can use to formulate our predictions. The marketing literature has conceptualized brand associations using both categorization theory and associative network theory. Categorization theory imposes a hierarchical structure on consumers' cognitive representation and emphasizes a brand's graded product category membership (its typicality) according to semantic feature similarities (Loken and Ward 1990; Sujan 1985). The focus of studies using the categorization model has been on examining product category–brand or brand–product relationships rather than brand–brand relationships.

In a brand portfolio, category membership is one of many ways that brands can be linked. Other ways that brands can be linked include the common parent brand name, phonetic similarity, spatial proximity (e.g., brands placed closely on the shelf), and goal congruence (Ratneshwar, Pechmann, and Shocker 1996). For example, Nescafé and Nestea may be linked in consumers' memory because of their phonetically similar brand names and because they share a similar trade dress and design, in addition to their common membership of the beverage category.

Associative network theory (Collins and Loftus 1975) provides a general representation that captures the many types of relatedness that can occur between brands in a portfolio. This theory conceptualizes brand knowledge as consisting of a brand node to which a variety of associations, such as brand claims, evaluations, and attributes, are linked (Keller 1993; Morrin 1999). An extension of this conceptualization of brands as nodes captures the relatedness between brands by positing linkages between nodes. Within a brand portfolio, firms' family-branding initiatives can be viewed as attempts to establish cognitive linkages between the parent brand and its subbrands or among subbrands. For example, Unilever's ice-cream brand in Europe, Ola, is associated with several subbrands that are often presented together in a single print advertisement. After they are established, these links can gain strength through repeated exposure and additional learning opportunities. Furthermore, because subbrands within a brand portfolio often share similar quality standards, brand image, and advertising execution elements, consumers may also use these additional cues to organize their mental representation of brand portfolios. Accordingly, we conceptualize the representation of brand portfolio information as a brand network that consists of a set of interlinked brand nodes. In a managerial illustration of brand architecture, Aaker and Joachimsthaler (2000) use specifications such as endorsed brands, subbrands, and cobrands to describe different brand entities involved in a brand portfolio. In this research, we use the specification "subbrand" to refer to any brand entity (e.g., Nescafé, KitKat) that carries a part of the product line within the parent brand portfolio (e.g., Nestlé).

Spillover as a Cognitive Process

Strength of Brand Relatedness and Spillover

Spillover effects in a brand network can be viewed as the combination of two consecutive processes: the retrieval of related nodes and their updating. Retrieval is considered the result of "spreading activation" through the associative network (Collins and Loftus 1975). In the context of brand portfolios, a brand node (referred to as the "origin brand") is primed and activated by external information, and this activation spreads to related brand nodes (referred to as "destination brands") through associative network linkages. Brand relatedness—the strength of association between brands—is reflected in the probability of brand retrieval and the level of activation at the destination brands (Morrin 1999; Nedungadi 1990). When strength of association is retrieved, it may also determine the extent of updating of these brands. Prior research in attitude and information integration (e.g., Olson and Zanna 1993) has suggested that brand evaluation updating occurs when consumers are exposed to valenced messages, and the extent of updating depends on message memory and the amount of elaboration triggered by the message. In the context of brand portfolios, strong associations between brands render the information at the origin brand more salient and relevant at the destination node (Chapman and Aylesworth 1999), and therefore a greater extent of updating in the destination node may be expected. The resultant updating of the destination node provides a measure of the spillover effect.

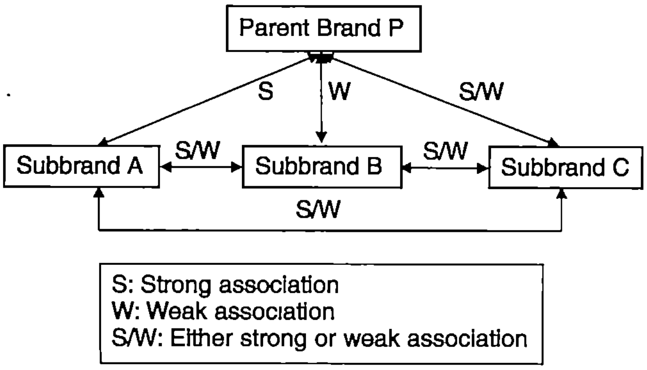
That the magnitude of spillover is a function of the strength of association has been shown in previous research. For example, Roehm and Tybout (2006) find that perceived similarity between brands makes a scandal about one brand more diagnostic to competitor brands and leads to more spillover. Their model measures similarity as a function of the shared attributes among brands. Therefore, similarity asymmetry and its consequence on spillover were not examined in their study.

Directionality of Brand Relatedness and Spillover

Linkages between cognitive entities may not be symmetric in strength because "the direction processed more frequently develops stronger relations" (Barsalou and Sewell 1985, p. 650). Marketing scholars who allude to directionality suggest that neglecting the directionality of brand associations can lead to erroneous predictions about consumers' brand-related behavior. For example, Farquhar and Herr (1993) and Herr, Farquhar, and Fazio (1996) suggest that brand-building activities focus on building and strengthening the directional association from the brand to associates, such as product category, product attributes, and usage situations, whereas brand-leveraging activities depend mainly on the directional associations from these other associates to the brand, leading Holden and Lutz (1992) to the exhortation, "Ask Not What the Brand Can Evoke; Ask What Can Evoke the Brand?"

Similarly, in a brand network, the directional strength of association between brand nodes may vary, and such differences can be critical in predicting the pattern of spillover in the network. In Figure 1, we illustrate the structure of a simple but typical brand portfolio with a parent brand, P (e.g., Kellogg), and three subbrands, A, B, and C (e.g., Special K). We first focus on the linkages between each of the subbrands and the parent brand. The parent brand–subbrand link reflects the strength of association from the parent brand to the subbrand and can be used to predict the extent to which the parent brand evokes the subbrand; conversely, the subbrand–parent brand link reflects the likelihood that the subbrand will evoke the parent brand.

FIGURE 1
An Example Associative Brand Network Representing a Brand Portfolio



If a distinction between the two directions of association exists, the directionality of association should influence the pattern of spillover. In other words, if the subbrand–parent brand link is not of the same strength as the parent brand–subbrand link, we expect to observe a different magnitude of spillover, depending on the direction. However, the question still remains, Which of the two links predicts the spillover on the parent brand when the crisis occurs at the subbrand, and vice versa? Given that negative events such as product-harm crises usually occur at the subbrand level, we consider the spillover of subbrand crises onto the parent brand. On the basis of the retrieval–updating process posited to underpin spillover, we propose that when a subbrand is the locus of the initial activation, the subbrand–parent brand link serves as the pathway for spreading activation and the consequent retrieval and updating of the parent brand node. Therefore, as a first step, we propose to test the extent to which spillover to the parent brand due to negative information at a subbrand is a function of the subbrand–parent brand link strength rather than the parent brand–subbrand link strength. Thus:

> H_1: Spillover from subbrand to parent brand is a function of the subbrand–parent brand link strength rather than the parent brand–subbrand link strength.

Asymmetry may also exist in the subbrand–subbrand link. Therefore, determining directionality of association is important in predicting spillover between subbrands. In other words, if the strength of association between subbrands is asymmetric, the magnitude of spillover will be asymmetric. Associative network theories indicate that the strength of a link depends not only on the quantity and quality of processing but also on the relative importance of the link to a conceptual node. Specifically, the amount of activation spreading from a brand node along a linkage is a function of the strength of that linkage relative to the sum of strengths of all linkages emanating from that brand node. Therefore, given a fixed level of activation at a node, the more links it radiates, the weaker is the activation that emanates toward any one link. This effect is referred to as the "fan" effect in Anderson's (1983) spreading activation model (see also Kahana 2002). In a brand network, not all subbrands have equal salience. Some subbrands are more salient or dominant in the network and have more links emanating from them than others. We expect this asymmetry in associations to lead to asymmetric spillover between subbrands. Thus:

> H_2: Spillover between subbrands is a function of the directional strength of association between subbrands. In particular,
>
> (a) The spillover between Subbrands A and B is asymmetric when the strength of association between A and B is asymmetric, and
>
> (b) Negative information at Subbrand A has greater spillover on Subbrand B than on Subbrand C when the A–B association is stronger than the A–C association.

Experiment 1a

The purpose of Experiment 1a is to examine the directionality and asymmetry of spillover in brand portfolios. A total of 387 students participated in a series of pretests and the main study in exchange for entry into lotteries with small prizes. The experimental stimuli we used in this study were snack bars and ice cream, products that are highly relevant and familiar to the participants. Of the 185 students who participated in the main study, 47.6% were female, 59.5% were business students, and 40.5% were law students. Participants in the pretests were drawn from the same pool as the main study.

Pretests

We conducted five pretests with the following objectives: (1) to select a suitable brand portfolio for hypothesis testing, (2) to test the fictitious product-harm crisis stories to be used as stimuli for readability and realism, and (3) to assess baseline familiarity and brand evaluations for brands in the selected portfolio. Pretest 1 measured the directional strength of association from subbrands to their parent brands using a computer-based response time sequential priming method. Response time sequential priming is well established in social psychology and consumer research as a method for measuring the strength of association between nodes in a cognitive structure (e.g., Bargh and Chartrand 2000; Herr, Farquhar, and Fazio 1996). Participants were exposed to a prime and then a target after a short delay (referred to as "stimulus onset asynchrony"). Following established practice (e.g., Herr, Farquhar, and Fazio 1996), we set the stimulus onset asynchrony at 750 milliseconds (ms). We asked right-handed participants to respond to the question of "relatedness" between two brands that appeared on a computer screen by pressing the "M" (yes) or "Z" (no) key (we used the reverse pattern of keys for left-handed participants) as rapidly as possible while remaining accurate in their responses. There were 34 participants in Pretest 1. The task consisted of four practice questions, followed by five blocks of eight questions each (four experimental questions and four filler questions) in a randomized order. The first word in each pair was a subbrand (e.g., Corn Flakes), and after a delay of 750 ms, a parent brand (e.g., Kellogg) that may or may not have been related appeared on the screen as the second word. The response latency, measured in ms, is taken as an inverse indicator of the strength of association between the two brands.[1] We tested several consumer packaged goods brand portfolios as part of the initial stimulus pool. Criteria for the selection of a stimulus set were that it needed to include several subbrands associated with the parent brand at varying levels of strength (see Figure 1). On the basis of the results of Pretest 1, we selected the Ola ice-cream brand portfolio (with subbrands Raket, Cornetto, and Magnum) for Experiment 1a. The results show that the strengths of the Magnum–Ola and Cornetto–Ola associations are both stronger than the Raket–Ola association (836 ms versus 1484 ms; $t_{28} = -2.71$, $p < .05$; 880 ms versus 1186 ms; $t_{27} = -2.18$, $p < .05$). Furthermore, Magnum and Cornetto have equivalent strengths of association to Ola (855 ms versus 902 ms; $t_{27} = -.38$, $p > .70$). Participants also evaluated six fictitious news-style stories with different product-harm scenarios. One was selected for its realism and readability. The selected story described a quality problem with one of the subbrands that causes consumers to fall ill.

Pretest 2 measured the directional strength of association from the parent brand to subbrands. There were 32 participants in Pretest 2. Using a method similar to that in Pretest 1, we found that for all three subbrands, the strengths of parent–subbrand (Ola–Magnum, Ola–Raket, and Ola–Cornetto) association were equivalent (Ola–Magnum versus Ola–Raket: 1187 ms versus 1324 ms; $t_{24} = -.98$, $p > .30$; Ola–Magnum versus Ola–Cornetto: 1195 ms versus 1272 ms; $t_{26} = -.93$, $p > .30$; Ola–Raket versus Ola–Cornetto: 1313 ms versus 1261 ms; $t_{28} = .45$, $p > .60$). This is in contrast to the inequivalence of the associative strengths in the other direction and provides initial corroboration that associative strength in brand networks is directional.

Pretests 3 and 4 measured the directional strength of association between the subbrands. Linkages between subbrands exist not only because of the common parent brand but also because of similar product features, complementary usage situations, and other commonalities. The reaction time measurement used in Pretests 1 and 2 is ideally suited to measuring single relationship recognition, such as filial relationship (e.g., between subbrands and the parent brand) or category membership. Because links between subbrands are multifaceted, we used a computer-assisted sequential priming task to measure "relatedness" on a seven-point scale (7 = "strongly related") (Chapman and Aylesworth 1999; Herr, Farquhar, and Fazio 1996). Before this task, participants read a passage describing how subbrands could be related in several ways, such as through common brand families, similar features, and similar usage occasions, with instructions to consider these types of linkages when making their judgments (Herr, Farquhar, and Fazio 1996). In Pretest 3, the first word in each pair of the priming task was a subbrand, followed by the appearance of another (related or unrelated) subbrand after a 750 ms delay. The remainder of the procedure was similar to Pretest 1. We switched the appearance sequence of the two subbrands in Pretest 4. There were 30 participants in Pretest 3 and 32 in Pretest 4. The results showed that the associations between Magnum and Raket and between Magnum and Cornetto are asymmetric in strength (Magnum–Raket versus Raket–Magnum: 2.10 versus 2.88; $t_{60} = -2.86$, $p < .01$; Magnum–Cornettó versus Cornetto–Magnum: 5.00 versus 5.70; $t_{60} = -2.30$, $p < .05$). There was no significant difference for the Cornetto–Raket pair (Cornetto–Raket versus Raket–Cornetto: 2.50 versus 2.84; $t_{60} = -1.29$, $p > .20$). Furthermore, Magnum is significantly more strongly associated with Cornetto than with Raket (5.00 versus 2.10; $t_{60} = 10.09$, $p < .001$), and Cornetto is significantly more strongly associated with Magnum than with Raket (5.70 versus 2.50; $t_{29} = 11.55$, $p < .001$). The Raket–Magnum and Raket–

---

[1] Participants who incorrectly answered the relatedness question apparently did not associate subbrands with the parent brand, and therefore we did not include them in further analysis. The same was done in Pretest 2.

Cornetto association is equivalent (2.88 versus 2.84; t_{31} = .14, $p > .80$). The relative strengths of association between each pair of brands in the Ola portfolio appear in Figure 2.

Pretest 5 measured baseline brand familiarity and brand attitudes among 74 participants. The results showed that participants were familiar with the Ola brands (M = 5.59, 5.91, 5.91, and 5.53 for Ola, Cornetto, Magnum, and Raket, respectively; 7 = "familiar"; $ps > .10$). Participants also had similar initial brand attitudes toward these brands (M = 5.62, 5.89, 5.53, and 5.52, respectively; 7 = "favorable"; $ps > .10$).

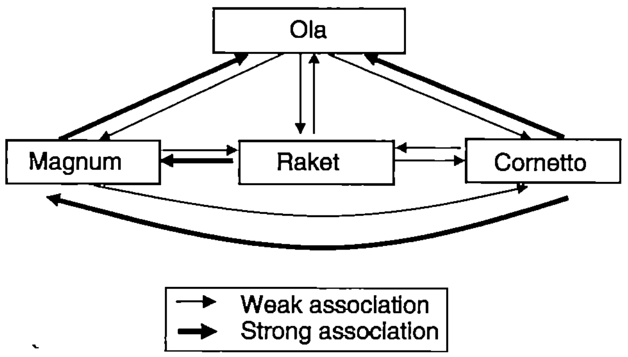
Experimental Procedure

We randomly assigned 185 participants to one of the three experimental groups and one control group. In the experimental groups, we employed negative information—in particular, product-harm crises—as an external information prime because the spillover effects through the brand network are likely to be pronounced and, therefore, more easily detectable from a methodological standpoint (e.g., Ahluwalia, Burnkrant, and Unnava 2000). Participants in each experimental group were exposed to the same crisis story about one of the three subbrands (Magnum, Cornetto, or Raket). In this manner, we manipulated the locus of the product-harm crisis to vary strength of association with other brands. After reading the crisis story, participants evaluated the three subbrands—Magnum, Cornetto, and Raket—in a counterbalanced order and the parent brand Ola. Next, participants evaluated the perceived severity of the crisis story on a seven-point scale ("not at all severe" and "very severe") and provided demographic information. Participants were then debriefed and thanked. Those in the control group evaluated the brands without reading the crisis story.

Dependent Variables

Consistent with previous research on brands and negative information, we employed a composite measurement of brand evaluations, including dimensions of brand attitude ("negative/positive," "bad/good," "unfavorable/favorable"), brand trust ("not at all trustworthy/very trustworthy," "not at all reliable/very reliable," "not at all dependable/very dependable"), perceived quality of the brand and the product ("low quality/high quality"), brand purchase likelihood ("not at all likely/very likely"), and brand desirability ("not at all desirable/very desirable") (Cronbach's α = .94) (Ahluwalia, Burnkrant, and Unnava 2000; Dawar and Pillutla 2000). The mean of the ten-item scale serves as the dependent variable.

FIGURE 2
The Relative Strengths of Association Between Each Pair of Brands in the Ola Portfolio



Results

The manipulation check showed that the crisis was perceived as equally severe in the three experimental groups ($ps > .10$). We found no systematic differences across gender or subject majors ($ps > .10$). H_1 predicted that spillover from subbrand to parent brand would be a positive function of the subbrand–parent brand link strength but not of the parent brand–subbrand link strength. As we indicated, the link strength from the parent to each of the subbrands is equivalent in our stimuli, but the subbrand–parent brand links vary in strength. We test whether crises with loci at strongly associated subbrands have a greater negative impact on parent brand evaluations than crises at weakly associated subbrands.

Pretest results showed that the Raket–Ola association is significantly weaker than both the Magnum–Ola and the Cornetto–Ola associations. The results of a one-way analysis of variance (ANOVA) with planned contrasts showed that crises at Magnum and Cornetto had significantly stronger effects on Ola than did a crisis at Raket ($\text{Ola}_{\text{Raket crisis}}$ = 4.66 versus $\text{Ola}_{\text{Magnum crisis}}$ = 4.13; t_{74} = 2.13, $p < .05$; $\text{Ola}_{\text{Raket crisis}}$ = 4.66 versus $\text{Ola}_{\text{Cornetto crisis}}$ = 3.69; t_{71} = 3.67, $p < .001$). Furthermore, we know that the strength of the Magnum–Ola association is equivalent to the strength of the Cornetto–Ola association. The results showed that the impact of the crisis at Cornetto on Ola was not significantly different from that of the crisis at Magnum (t_{91} = –1.67, p = .10). Therefore, H_1 is supported; the magnitude of spillover at the parent brand is positively related to the strength of subbrand–parent brand association. Means of brand evaluations in the control and experimental groups appear in Table 1.

H_2 predicted that spillover between subbrands would be a function of the directional strength of association between them. H_{2a} tests whether the spillover between Subbrands A and B is asymmetric when the strength of association between them is asymmetric. In line with the pretests, the Raket–Magnum association is stronger than the Magnum–Raket association. The results of a 2 × 2 between-subjects ANOVA with one factor for Crisis Locus (Magnum versus Raket) and the other factor representing Crisis (crisis group versus control group) showed that the Raket crisis had a greater spillover on Magnum evaluations ($\text{Magnum}_{\text{control}}$ = 5.65 versus $\text{Magnum}_{\text{experiment}}$ = 4.91; t_{79} = 3.87, $p < .001$) than did the Magnum crisis on Raket evaluations ($\text{Raket}_{\text{control}}$ = 5.06 versus $\text{Raket}_{\text{experiment}}$ = 4.99; t_{91} = .36, $p > .70$; $F(1, 170) = 5.46$, $p < .05$). Next, we also know that the Cornetto–Magnum association is stronger than the Magnum–Cornetto association. The results showed that the Cornetto crisis had a greater spillover on Magnum evaluations ($\text{Magnum}_{\text{control}}$ = 5.65 versus $\text{Magnum}_{\text{experiment}}$ = 3.80; t_{84} = 7.89, $p < .001$) than did the Magnum crisis on

TABLE 1
Mean Differences (Standard Error) of Brand Evaluations in Experiment 1a

Crisis Locus	Brand Evaluation in Control and Different Crisis Groups: Raket	Magnum	Cornetto	Ola
Raket	2.64** (.1893)	4.91** (.1241)	4.70* (.1412)	4.66** (.1693)
Magnum	4.99 (.1297)	3.02** (.1965)	4.24** (.1537)	4.13** (.1854)
Cornetto	4.70 (.1813)	3.80** (.1827)	2.44** (.1550)	3.69** (.1837)
Control	5.06 (.1659)	5.65 (.1473)	5.34 (.1209)	5.83 (.1762)

*$p < .01$ (means in the crisis group compare with the mean in the control group).
**$p < .001$ (means in the crisis group compare with the mean in the control group).

Cornetto evaluations ($\text{Cornetto}_{\text{control}} = 5.34$ versus $\text{Cornetto}_{\text{experiment}} = 4.24$; $t_{91} = 5.64$, $p < .001$; $F(1, 176) = 5.41$, $p < .05$).

H_{2b} tests whether the crisis at Subbrand A has a greater spillover onto B than C when the A–B association is stronger than the A–C association. A 2 × 2 mixed ANOVA with a within-subjects factor for strength of association (strong versus weak; the same participant evaluated both strong and weak subbrands associated with the subbrand implicated in the crisis) and a between-subjects factor representing a crisis (crisis group versus control group) showed that the crisis at Magnum had greater spillover on the strongly associated Cornetto than on the weakly associated Raket ($\text{Cornetto}_{\text{control}} = 5.34$ versus $\text{Cornetto}_{\text{experiment}} = 4.24$; $t_{91} = 5.64$, $p < .001$; $\text{Raket}_{\text{control}} = 5.06$ versus $\text{Raket}_{\text{experiment}}$ 4.99; $t_{91} = .36$, $p > .70$; $F(1, 91) = 15.62$, $p < .001$). Similarly, we also found that the crisis at Cornetto had a greater spillover on the strongly associated Magnum than on the weakly associated Raket ($\text{Magnum}_{\text{control}} = 5.65$ versus $\text{Magnum}_{\text{experiment}} = 3.80$; $t_{84} = 7.89$, $p < .001$; $\text{Raket}_{\text{control}} = 5.06$ versus $\text{Raket}_{\text{experiment}} = 4.70$; $t_{85} = 1.43$, $p > .10$; $F(1, 85) = 25.73$, $p < .001$). Overall, the results support H_2.

Discussion

In Experiment 1a, we measured the directional strength of association between subbrands and the parent brand and between subbrands and found that asymmetric strength of association predicts asymmetric spillover between brands. These results indicate that both strength and directionality of association are significant predictors of spillover effects of a product-harm crisis, both from subbrands to the parent brand and between subbrands. However, a potential limitation of Experiment 1a is that we measured rather than manipulated the strength of association. Although this provides a valuable test of the hypotheses in a "naturalistic" setting, the evidence presented is correlational, not causal. In addition, the specificities of stimuli employed could potentially confound the result. We conducted Experiment 1b to address these issues.

Experiment 1b

Experimental Procedure

We used the Nestlé portfolio with the subbrands Nuts and Crunch. For simplicity, we focused on the subbrand–parent brand links in this experiment. In memory research, the rehearsal or repeated-exposure task is commonly used to strengthen the association between concepts (e.g., Posavac et al. 2001). In the experimental group, participants performed a computer-assisted priming task in which the Nuts–Nestlé and Crunch–Nestlé associations (and several filler brand associations) were strengthened through repeated sequential exposure. Each repetition required participants to respond to a question about the "relatedness" of the two brands, and an incorrect response triggered a feedback message. On the basis of previous studies (e.g., Gürhan-Canli and Maheswaran 1998), we minimized the possibility of demand effects from the repeated-exposure task by providing a cover story and instructions that focused participants on the task, and we minimized the cognitive resources available to speculate about other possible purposes of the task by including filler brands in the repeated-exposure task to minimize the possibility of making the stimulus brands salient and by scheduling a break after the completion of the priming task to clear participants' short-term memory. After the break, participants completed a questionnaire with a crisis (the same crisis used in Experiment 1a) at the Nuts brand. The remainder of the procedure was similar to Experiment 1a. Participants in the control group completed the same procedure as those in the experimental group, but we did not manipulate the strength of association between Nuts and Nestlé and between Crunch and Nestlé through repeated exposure.

Results

Manipulation checks showed that the Nuts–Nestlé association was significantly stronger in the experimental group than in the control group (413 ms versus 1240 ms; $t_{22} = -4.75$, $p < .001$), as was the strength of the Crunch–Nestlé association (470 ms versus 1258 ms; $t68 = -7.32$, $p < .001$). The results showed that compared with the control group, the decline in the brand evaluations of the parent brand, Nestlé, due to the crisis spillover was greater when the strength of association was experimentally increased through rehearsal ($\text{Nestlé}_{\text{control}} = 4.29$ versus $\text{Nestlé}_{\text{experiment}} = 3.25$; $t_{45} = 2.29$, $p < .05$). Notably, the inclusion of the second subbrand, Crunch, enables us to gauge the spillover across subbrands. The results showed that the Crunch subbrand was more severely affected by the crisis at the Nuts subbrand in the experimental group than in the control group ($\text{Crunch}_{\text{control}} = 4.42$ versus

$\text{Crunch}_{\text{experiment}} = 3.29$; $t_{40} = 3.11$, $p < .01$). In line with associative network theory, activation at a brand node spreads out along the paths to connected nodes, and the stronger the activation, the farther the activation spreads along the paths. Therefore, we surmise that the activation at Nuts spreads to Crunch through the Nuts–Nestlé–Crunch paths and that the increase in the Nuts–Nestlé association increased the Nuts–Crunch spillover.

Discussion

These results indicate that the manipulated increase in the strength of the subbrand–parent brand association leads to greater crisis spillover from the subbrand to the parent brand. This finding increases our confidence in the results of Experiment 1a about the causal relationship between the directional strength of association and the magnitude of spillover effects. Overall, the results of Experiment 1 showed that associations between brands can be asymmetric (e.g., associations between Raket and Magnum and between Cornetto and Magnum) or symmetric (e.g., association between Raket and Cornetto),[2] and asymmetric associations lead to asymmetric spillover effects. Given this finding, the question of what determines the directionality of association strength becomes relevant. We conducted Experiment 2 to delve into the antecedents of directionality.

Experiment 2a

Directionality of Brand Relatedness

Two theoretical explanations exist for the directionality of brand associations. Both suggest that asymmetry is a function of the number of associations of a brand. However, they differ when it comes to making predictions about asymmetry as a function of the salience of associations. Experiment 2a tests their common prediction, and Experiment 2b examines their divergent predictions. As we discussed previously, the first theory is the fan effect from associative network theory (Anderson 1983; Collins and Loftus 1975), which suggests that given a fixed level of activation, the more links a node has, the less activation is available for any one link. When two related brands have different numbers of associations, the strength of the link from the brand with fewer associations to the brand with more associations is stronger than the reciprocal link.

The second theoretical explanation resides in Tversky's (1977) contrast model. This model suggests that the perceived similarity between Brand A and Brand B is expressed as a function of their common features (A ∩ B) and distinctive features (A – B: distinctive features of Brand A; B – A: distinctive features of Brand B):

$$S(a, b) = \theta f(A \cap B) - af(A - B) - \beta f(B - A).$$

The parameters θ, a, and β are weights that represent the importance of each feature set for the similarity judgment, and the f function measures the salience of features. When judging the similarity of Brand A to Brand B, Brand A's features are weighed more heavily than brand B's ($a > \beta$). Thus, the distinctive features of Brand A detract more from similarity than the distinctive features of Brand B. Therefore, when Brand A (e.g., Shasta Cola) has fewer distinctive features (and, therefore, fewer features in total) than Brand B (e.g., Coca-Cola) (assuming that features are equally salient), this model predicts that Brand A will be judged more similar to Brand B than will Brand B to Brand A. In a brand portfolio, subbrands may be linked to different numbers of features because of different levels of brand familiarity, order of entry to the marketplace, or brand positioning. On the basis of the preceding discussion, we expect that the subbrand with fewer features will be perceived as more similar to the subbrand with more features than the reciprocal.[3] The directional similarity judgment (e.g., similarity of Brand A to Brand B) is said to be positively related to the level of activation spread from one brand to the other—in other words, the directional strength of association from Brand A to Brand B (Ulhaque and Bahn 1992). Therefore, both associative network theory and the contrast model predict the following:

> H_3: The strength of association from the brand with fewer associations to the brand with more associations is stronger than the reciprocal strength of association.

Stimulus Development

Two fictitious subbrands with different numbers of brand associations were required to test H_3. First, we adapted six fictitious subbrand names (e.g., Tristix, Portello, Wheatello) from existing brand names in the product categories of ice cream, candy bars, and cookies; we told participants that these subbrands were from parent brands such as Good Humor, Nestlé, and Kraft. To avoid the confounding effect of brand specifics, we invited 12 participants to test the memorability and suitability of these fictitious brand names (Schmitt, Tavassoli, and Millard 1993). The results showed no differences in memorability and suitability among them (*ps* > .10). Participants evaluated how realistic the associations presented for each fictitious brand were. The results showed that the descriptions of the Tristix and Portello ice-cream products (developed for the Good Humor portfolio) were considered the most realistic. Second, we selected nine associations (eight distinctive and one common) for the Tristix subbrand and three (two distinctive and one common) for Portello to manipulate the number of associations linked to each brand. These associations describe the product category and company that the subbrands belong to, the type of ice cream, packaging, flavor, ingredients, nutrition, and so forth.

---

[2] The results showed that the spillover of the Raket crisis on Cornetto evaluations ($\text{Cornetto}_{\text{control}} = 5.34$ versus $\text{Cornetto}_{\text{experiment}} = 4.70$; $t_{79} = 3.41$, $p < .01$) was not significantly different from the spillover of the Cornett crisis on Raket evaluations ($\text{Raket}_{\text{control}} = 5.06$ versus $\text{Cornetto}_{\text{experiment}} = 4.70$; $t_{85} = 1.43$, $p > .10$; $F(1, 164) = .61$, $p > .40$).

[3] According to Keller (1993), brand features are also associations.

Experimental Procedure

There were 48 participants in the experiment, and they were broken down into groups of up to 15 participants per session in a behavioral research laboratory. We instructed participants to learn different numbers of associations of the fictitious brands Tristix (nine associations) and Portello (three associations) and the filler brand Activia (three associations). To emphasize each association and avoid confusion, we presented these associations on 15 cards and separated them into three piles, one pile for each brand. We stacked the three piles of cards in a counterbalanced order to avoid the effect of directional learning. We instructed participants to remember as much information as possible in five minutes and provided them with paper and a pencil to record the associations, thus facilitating learning. Next, participants completed an unrelated filler questionnaire consisting of one page of text and 20 questions to clear their short-term memory. Then, we tested the directional strength of association between the brands Tristix and Portello together with 32 filler brands in a sequential priming task (the same method as in Pretests 3 and 4 of Experiment 1a). We tested each participant randomly on either the forward or the reciprocal link. Immediately after this task, participants wrote down the features they could recall for the brands Tristix and Portello. Finally, we asked participants about their familiarity with ice-cream products and then debriefed them.

Results

Two judges were invited to count the number of recalled features for Tristix and Portello. The manipulation check showed that Tristix was linked to significantly more associations in participants' memory than Portello (6.05 versus 2.85; $t_{47} = 15.48$, $p < .001$). Furthermore, the results showed that there was no significant difference in category familiarity between the group tested on the forward strength of association and the group tested on the reciprocal strength of association (4.44 versus 4.80; $t_{46} = 1.02$, $p > .30$).

H_3 predicted that the directionality of association strength would be influenced by the number of associations linked to brands. The results of an independent t-test showed that the Portello–Tristix association was significantly stronger than the Tristix–Portello association (5.00 versus 4.12; $t_{46} = 2.33$, $p < .05$). This result supports H_3; the associative link from brands with fewer associations (Portello) to brands with more associations (Tristix) was stronger than the reciprocal link.

Discussion

In Experiment 2a, we experimentally manipulated the number of associations linked to two fictitious subbrands and found that the brand with fewer associations was linked more strongly to the brand with more associations than the reciprocal link. This result supported the predictions of both associative network theory and the contrast model about the effect of the number of associations (or features) on the directional strength of association. Together with the findings in Experiment 1, we showed that asymmetric strength of association leads to asymmetric spillover between brands and that the number of associations linked to each brand is an important antecedent of association asymmetry.

However, the strength of directional association is influenced not only by the structural properties of brand associations, such as the number of associations linked to each brand, but also, potentially, by the salience of these associations. This contextual property (i.e., association salience) is managerially important because when a brand is affected by negative news, remedy strategies often involve changing the salience of certain brand associations. For example, managers may attempt to quarantine the affected brand by differentiating the affected brand (emphasizing its unique features) from the other brands in the portfolio. From a theoretical perspective, the effectiveness of such strategies depends on whether the priming of brand associations influences directional strength of association from the affected brand to other brands. Experiment 2b examines these managerial and theoretical issues.

Experiment 2b also allows us to overcome some of the limitations of Experiment 2a. For example, because Experiment 2a employed fictitious brand stimuli that participants first encountered in the laboratory, a common association between the brands may have been that their associations were learned in the same laboratory setting. To overcome this potential shortcoming, we employed real brand stimuli in Experiment 2b.

Experiment 2b

Impact of Priming Brand Associations

Intuitively, actively differentiating the harmed brand by, for example, proclaiming its distinctive features and muting its common features may limit the extent and scope of spillover because it reduces its strength of association with other brands. However, the action of differentiation may also prime the harmed brand further. Associative network theory and the contrast model make divergent predictions.

First, the contrast model suggests that the perception of similarity between brands increases with the measure of common features and decreases with the measure of distinctive features. The contribution of these feature sets to overall similarity depends on both the weights (θ, a, and β) and the salience (the f function) of the features. The weights of features depend on the task context (e.g., similarity or dissimilarity judgment) and the direction of judgment (e.g., features of Brand A weigh more heavily than those of Brand B when judging similarity of Brand A to Brand B). The salience of a stimulus depends on several factors, including intensity, familiarity, recency, frequency, and informational content (Tversky and Gati 1978). Given the fixed weight of each feature set, priming the distinctive features of a brand will increase the salience of these features and therefore should decrease the perceived similarity of this brand to other brands. Conversely, if the common features are primed, the salience of the common feature set will be increased, heightening the perceived similarity of the brands.

Second, associative network theory suggests that the activation of a destination node depends not only on the number of links emanating from the origin node but also on the strength of activation of the origin node. The strength of activation may vary as a function of the salience of brand priming (Nedungadi 1990), and in turn brand salience may be heightened when an association (e.g., a feature) of the origin brand is primed. In other words, priming associations of an origin brand, regardless of whether they are distinctive or common associations, should increase the level of activation at the origin brand and thus increase the level of activation at the destination brand. Therefore, we propose the following competing hypotheses:

> H_4: Under the contrast model, priming common associations of an origin brand increases the strength of association from the origin brand to the destination brand, but priming distinctive associations of the origin brand decreases the strength of association from the origin brand to the destination brand.
>
> H_5: Under associative network theory, priming associations of an origin brand, regardless of whether they are distinctive or common associations, increases the strength of association from the origin brand to the destination brand.

Stimulus Development

To test these hypotheses, we require two suitable subbrands that have distinctive and common associations and are linked to different numbers of associations. We presented six subbrands (e.g., Nescafé, KitKat, Corn Flakes) from two brand portfolios (Nestlé and Kellogg) in a randomized order to 14 participants. Participants were asked to "list the features of these brands that come to mind" and were given two minutes for each brand (Johnson 1986). Two judges counted the number of features. The results showed that among the six brands, Nescafé had significantly fewer features than Nestea (2.28 versus 5.36; $t_{13} = -5.40$, $p < .001$). Next, the two judges coded the common and distinctive features listed for the Nescafé and Nestea brands. Features were coded as "unique associations" if different wording was used, and they were coded as "common associations" if similar wording was used (Johnson 1986). The judges were in agreement 90% of the time, and disagreements were resolved by discussion. The listed common associations included the parent brand, Nestlé, and the beverage product category (e.g., "It is a drink"). For the distinctive associations, the most frequently mentioned ones were that Nestea is an iced tea drink and had different flavors, but Nescafé is an instant coffee and is sold in the form of powder (e.g., "ground coffee").

Experimental Procedure

We randomly assigned 132 participants to two control groups and four experimental groups in a 3 (no priming, distinctive association priming, common association priming) × 2 (direction of association) between-subjects design. In the experimental groups, participants were first instructed to state their agreement ("agree" or "disagree") with the information on six cards; each card described a brand association. Of these six statements, four were filler statements, and two statements primed two common or two distinctive associations of the stimulus brand. Next, we tested the directional strength of association between the stimulus brands together with filler brands in a sequential priming task (the same method as in Pretests 3 and 4 of Experiment 1a). After this task, we gave participants an unrelated filler questionnaire to clear their short-term memory. Finally, participants listed as many features of Nescafé and Nestea as they could recall in two minutes for each brand and then were debriefed.

Results

The manipulation check showed that Nestea was linked to significantly more associations in participants' memory than Nescafé (5.05 versus 2.15; $t_{131} = 19.56$, $p < .001$). The results in the control group showed that the Nescafé–Nestea association is significantly stronger than the Nestea–Nescafé association (3.52 versus 2.52; $t_{46} = 2.48$, $p < .05$). This result further corroborates the finding of Experiment 2a with real brand stimuli. Specifically, the strength of association is asymmetric when consumers associate different numbers of associations with two brands.

H_4 and H_5 offered different predictions about the impact of priming brand associations on the directional strength of association. Specifically, H_4 predicted that the linkage would be strengthened when priming common associations and weakened when priming distinctive associations, whereas H_5 predicted an equivalent effect of priming common and distinctive brand associations. We first analyzed the impact of priming common associations. The results showed that the Nescafé–Nestea association was significantly strengthened when the common associations of Nescafé were primed (3.52 versus 4.55; $t_{43} = -2.05$, $p < .05$), as was the strength of the Nestea–Nescafé association when Nestea was primed (2.52 versus 3.84; $t_{42} = -2.99$, $p < .01$).

Next, we analyzed the impact of priming distinctive associations. The results showed that the Nescafé–Nestea association increased when the distinctive associations of Nescafé were primed (3.52 versus 4.30; $t_{44} = -1.71$, $p < .05$), as did the Nestea–Nescafé association when Nestea was primed (2.52 versus 3.70; $t_{43} = -2.37$, $p < .05$). These findings support H_5; priming both common and distinctive brand associations increases the directional strength of association.

Discussion

In Experiment 2b, we found that compared with the control group, in which no association was primed, priming common and distinctive associations both increased the directional strength of association between brands. This finding favors the prediction of associative network theory over that of the contrast model. In line with associative network theory, the findings show that priming associations increases the level of activation emanating from the origin brand. The increased activation at the origin brand can temporarily increase the activation at the destination brand, which affects the strength of relatedness between brands. In addition, our results for the asymmetric association between the real subbrands Nescafé and Nestea provide further evidence in support of H_3.

Overall, Experiments 1 and 2 examined the directionality of brand association and its impact on spillover in a brand portfolio. In Experiment 1, we found that asymmetry exists and governs spillover. In Experiment 2, we found that the directional strength of association is influenced by both the number of associations linked to each brand and the salience of those associations.

General Discussion

Cognitive brand relatedness makes brands subject to spillover effects from other brands in a brand portfolio. These indirect antecedents of brand evaluations have received relatively little attention in the marketing literature. Our research examines the negative spillover of product-harm crises in brand portfolios. We found that the pattern of spillover is a function of both strength and directionality of brand association. In turn, the strength and the directionality of brand linkages are influenced by the number of associations linked to each brand and the salience of these associations.

Theoretical Implications

Our findings carry implications for the cognitive representation of brand portfolios and the consequences of the properties of these representations on spillover effects. Previous research (e.g., Nedungadi 1990) that examined indirect effects between brands through product category linkages suggests that there are asymmetric effects between brands when they are associated with the category at varying levels of strength. However, we found that the magnitude of spillover is predicted by the total strength of association between subbrands rather than by the strength of filial linkages between subbrands and the parent brand. For example, Cornetto has stronger associations with Ola than does Raket, but the total strength of the Cornetto–Raket association is equivalent to that of the Raket–Cornetto association, which results in symmetric spillover between Raket and Cornetto. Similarly, both Magnum and Cornetto are strongly associated with Ola, but the total strength of the Magnum–Cornetto association is weaker than that of the Cornetto–Magnum association, which leads to asymmetric spillover between the two subbrands. These findings suggest that many types of commonality linkages define relationships between brands in a portfolio, and this can help predict spillover.

The importance of directionality of linkages for marketing phenomena has been conceptually highlighted in previous research, but it has never been empirically tested (Farquhar and Herr 1993; Holden and Lutz 1992). Assuming that linkages are nondirectional, as previous studies have done, leads to the measurement of one of the directions of the linkage and ignores the other. This omission may lead to erroneous predictions if the incorrect direction of the linkage is measured in predicting spillover. Our studies not only provide empirical tests of the directionality hypothesis but also examine the consequences of such directional linkages on spillover in brand portfolios. We find that accurate predictions of spillover cannot be made without specifying the direction of the linkage between two brands. Spillover from Brand A to Brand B depends on the directional strength of association from Brand A to Brand B and not that from Brand B to Brand A.

The finding that asymmetric associations lead to asymmetric spillover also has implications for the interpretation of the spillover effects found in previous studies. For example, Roehm and Tybout (2006) find that a brand scandal may affect competitor brands that are perceived as similar to the affected brand. Competing brands may preemptively issue denials in an attempt to distance themselves from the source of the scandal. However, issuing a denial is effective only when scandal spillover has indeed occurred. Our findings suggest that a more complex set of effects may be at play. For example, perceptions of similarity between a brand and its competitor brands may be asymmetric, leading to the possibility of asymmetric spillover. In other words, given the same scandal, a brand may not affect its competitor brand to the same extent as a scandal for the competitor would affect the original brand. Therefore, issuing a denial should not be an automatic reaction when a related brand in a category is implicated in a scandal.

To understand the sources of asymmetric linkage better, we examined factors that influence the directional strength of association between brands. We found that brands with fewer associations have stronger linkages to brands with more associations than vice versa. This finding provides empirical evidence for both the "fan effect" in associative network theory and the "asymmetric similarity judgment" in the contrast model. We also found that priming brand associations, regardless of whether they are common or distinctive associations, can increase the directional strength of association. On the one hand, this finding supports associative network theory in that the increase in the level of activation at the origin brand can strengthen its association with the destination brand. On the other hand, it may also corroborate some of the previous discussions on the limitations of the contrast model. For example, Johnson (1986) suggests that despite its generality, the contrast model does not specifically address the psychological process of judging similarity or how the contrasting of features proceeds. Therefore, simply priming the distinctive features of a brand may not necessarily lead to brand differentiation. For example, consumers' processing goals may influence the judgment of similarity. Consumers who are motivated to differentiate two brands are more alert to dissimilarities than consumers who just look for general information.

Finally, the findings of association asymmetry between brands in a brand portfolio also help us understand association asymmetry between other types of cognitive entities in consumers' memory. For example, previous research (e.g., Herr, Farquhar, and Fazio 1996) indicates that the strength of the category–brand association (referred to as "category dominance") may be asymmetric to that of the brand–category association (referred to as "instance dominance"). However, it is unclear when and why this asymmetry occurs. On the basis of our findings, we surmise that the brand–category asymmetry may occur when consumers link different numbers of associations to the brand and to the category information node in their memory. It is also possible that a brand–category linkage will become more category dominant when category-related information is primed, and it will become more instance dominant when brand-related information is primed.

Managerial Implications

In recent years, many consumer goods companies, including Nestlé, Unilever, and S.C. Johnson, have pursued an umbrella-branding strategy, placing a parent brand on product packs and advertisements of their subbrands. This deliberate attempt to increase the strength of association between subbrands and a parent brand is driven by the potential benefits of marketing efficiency. However, our results suggest that these need to be weighed against the risks of negative spillover when one subbrand suffers an adverse event. As part of a portfolio, brands are vulnerable not just to the risks inherent in their own products and actions but also to those of other brands in the portfolio. The likelihood of adverse events that can have spillover effects may be low, but the downside is dire and exacerbated by the interlinkages between brands.

The directionality of association strength and its effect on spillover carries implications for the management of negative spillover effect. For example, a crisis at Subbrand A may not influence evaluations of Subbrand B to the same extent as the same crisis at Subbrand B influences evaluations of Subbrand A. Managers need to be aware of this asymmetric spillover to design efficient remedy strategies that avoid exposing brands to unintended risk. Furthermore, if risks of adverse events are relatively high, as they may be in some product categories, such as pharmaceuticals or children's toys, but the benefits of parent branding are still attractive, the architecture of a brand portfolio should be designed to take advantage of associative asymmetry. Brand architectures could be designed to limit the potential spillover from subbrands considered at risk by having weaker associations with the parent brand while still maintaining potential benefits of positive spillover in the reciprocal direction.

Lured by the effect of association asymmetry, we examined the antecedents of association asymmetry and found that the number of associations linked to each brand was a significant factor. This finding has implications for brand positioning in a brand portfolio. For example, when a firm introduces a new brand, if the goal is to position the brand closely to an established brand, it may be wise to limit the number of associations introduced for the new brand. Conversely, if the goal is to differentiate the new brand, more associations will effectively help reduce the perceived similarity of this brand to other brands in the portfolio.

Furthermore, we found that priming brand associations can, at least temporarily, increase the directional strength of association from the brand to related brands. This finding indicates that when a subbrand is affected by negative information, it may not be wise to prime the distinctive associations of that brand in the hope of limiting spillover. Instead, advertising emphasis should shift to underscore the positive equity of the parent brand and/or other subbrands while advertising of the affected brand is halted. If firms do this, the positive associations of other brands in the portfolio can help neutralize the negative evaluations of the subbrand while minimizing the negative reciprocal spillover from the subbrand to the parent brand and/or other subbrands.

The finding that priming increases the directional strength of association also helps explain why negative information can sometimes affect brands that are considered only weakly related to the affected brand. Specifically, when a brand is implicated in a negative message (e.g., a product-harm crisis), it is also primed in consumers' memory. This priming may temporarily increase the directional strength association from the affected brand to the previously weakly related brand, making it a potential target for contamination. To examine this possibility, we conducted a posttest (n = 45) to examine the impact of priming brands in negative information on the direction strength of association. The results showed that compared with the control group, respondents associated the subbrand Nestea significantly more strongly with the fictitious subbrand Nescuit (a chocolate milk cookie from Nestlé) when Nestea was implicated in a fictitious negative news report (3.06 versus 4.14; $t_{43} = 2.10$, $p < .05$). This result indicates that the scope of contamination imposed by a negative message in a brand portfolio may be larger than what is normally expected. Managers need not only to prepare remedies actively for brands that are obviously related to the affected brand but also to keep a close eye on brands that seem to be less related and safe.

Finally, although we chose a particular type of brand portfolio in this research to test our hypotheses, our findings carry implications for other types of brand portfolio configurations as well. Aaker and Joachimsthaler's (2000) brand relationship spectrum depicts interbrand relationships with four types of brand portfolios: house of brands, endorsed brands, subbrands, and branded house. In the house-of-brands configuration, brands in a portfolio are often independent from one another. In this condition, if there is no cognitive relatedness between brands, no spillover should occur. In the endorsed-brands and subbrands configurations, consumers are often aware of the linkage between brands. Therefore, our predictions about the causal relationship between directional strength of association and spillover apply most explicitly in these two conditions. In the branded-house configuration, all products share the same brand name. In this condition, the brand and the product are often considered one conceptual node because priming the product automatically primes the parent brand name. In other words, in this type of portfolio, the focus is on the interproduct linkages (e.g., Balachander and Ghose 2003) rather than on interbrand linkages.

Limitations and Future Research Directions

In this research, we controlled for or kept constant several factors that might potentially moderate the relationship between our independent and dependent variables. In this section, we describe some of these because they provide avenues for further research.

We predicted the pattern and magnitude of spillover on the basis of the strength and directionality of associative linkages between brands in a portfolio. However, in examining the effects of external information on a brand network, we found that the characteristics of the information itself can also potentially influence outcomes. For example, previous research has shown that severe crises and crises that are perceived as avoidable cause more feelings of anger and stronger emotional reactions toward the harmed brand (Folkes 1984; Weiner 2000). Next, we investigated the impact of primary network features, strength, and directionality of association on the spillover effect. However, other features of a network, such as the strength of the node itself (e.g., a major versus minor brand) and the density of a network (e.g., the number of brands in a portfolio), may also influence the nature and extent of spillover. Further research on factors such as information characteristics and other network features that might moderate negative information processing would provide a more complete understanding of the spillover effects in a brand network.

We employed negatively valenced information as a stimulus to examine how its effects spread through the network. However, an open question is whether our findings can be equally applied to the spillover of positive information. For example, negative information has been shown to be more influential than positive information in forming evaluations and decision making (Skowronski and Carlston 1989), but previous research has also indicated that the greater influence of negative information (versus positive information) is not due to its negative valence but rather to its statistically rare, distinctive, and nonmodal characteristics (Scott and Tybout 1981). Therefore, we would expect a similar pattern of spillover of negative versus positive information if both types of information were similarly accessible, diagnostic, and statistically rare. Further research might verify this prediction.

We examined the effect of the number and salience of associations linked to each brand on the association asymmetry between brands. However, literature indicates that other factors may also potentially account for asymmetry in association between pairs of objects. One such factor is the direction of processing (Barsalou and Sewell 1985). Kahana (2002) also indicates that the direction of learning two words may lead to association asymmetry between the two words (e.g., river–bank). In this research, we controlled for this factor by counterbalancing the presentation order of fictitious brands in Experiment 2a and by randomly assigning subjects to different groups in Experiment 2b in which real brands were used. Further research might experimentally manipulate the learning sequence of brands to examine the effect of processing direction on association asymmetry.

Finally, in a brand portfolio, brands may be linked by semantic similarities, such as common attributes or a shared product category (e.g., KitKat–Smarties), but they may also be linked through more superficial lexical similarities, such as phonetically or orthographically similar brand names (e.g., Nescafé–Nestea). Brand linkages based on these two types of relationships (semantic versus lexical) may have different levels of strength, depending on the way brand information was learned. For example, when brand information is learned intentionally, consumers tend to rely on semantic relationships to establish links between brands. However, when brand information is learned incidentally, consumers tend to rely on heuristics (e.g., lexical cues) to associate brands. Further studies might investigate the relative strength of these two types of relationships under different learning environment.

REFERENCES

Aaker, David A. and Erich Joachimsthaler (2000), "The Brand Relationship Spectrum: The Key to the Brand Architecture Challenge," *California Management Review*, 42 (Summer), 8–23.

Ahluwalia, Rohini, Robert E. Burnkrant, and H. Rao Unnava (2000), "Consumer Response to Negative Publicity: The Moderating Role of Commitment," *Journal of Marketing Research*, 37 (May), 203–215.

Anderson, John R. (1983), "A Spreading Activation Theory of Memory," *Journal of Verbal Learning and Verbal Behavior*, 22 (3), 261–95.

Balachander, Subramanian and Sanjoy Ghose (2003), "Reciprocal Spillover Effects: A Strategic Benefit of Brand Extensions," *Journal of Marketing*, 67 (January), 4–14.

Bargh, John A. and Tanya L. Chartrand (2000), "Studying the Mind in the Middle: A Practical Guide to Priming and Automaticity Research," in *Handbook of Research Methods in Social Psychology*, H. Reis and C. Judd, eds. New York: Cambridge University Press, 253–85.

Barsalou, Lawrence and Daniel Sewell (1985), "Contrasting the Representation of Scripts and Categories," *Journal of Memory and Language*, 24 (6), 646–65.

Chapman, Kenneth J. and Andrew Aylesworth (1999), "Riding the Coat-Tails of a Positive Review: Rave Reviews and Attitude Transfer," *Journal of Consumer Marketing*, 16 (4–5), 418–41.

Collins, Allan M. and Elizabeth F. Loftus (1975), "A Spreading-Activation Theory of Semantic Processing," *Psychological Bulletin*, 82 (6), 407–428.

Dawar, Niraj and Madan M. Pillutla (2000), "Impact of Product-Harm Crises on Brand Equity: The Moderating Role of Consumer Expectations," *Journal of Marketing Research*, 37 (May), 215–26.

Farquhar, Peter H. and Paul M. Herr (1993), "The Dual Structure of Brand Associations," in *Brand Equity and Advertising: Advertising's Role in Building Strong Brands*, David A. Aaker and Alexander L. Biel, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 263–77.

———, ———, and Russell H. Fazio (1990), "A Relational Model for Category Extensions of Brands," in *Advances in Consumer Research*, Vol. 17, M.E. Goldberg, G. Gorn, and R.W. Pollay, eds. Provo, UT: Association for Consumer Research, 856–60.

Folkes, Valerie S. (1984), "Consumer Reactions to Product Failure: An Attributional Approach," *Journal of Consumer Research*, 10 (March), 398–409.

Gürhan-Canli, Zeynep and Durairaj Maheswaran (1998), "The Effects of Extensions on Brand Name Dilution and Enhancement," *Journal of Marketing Research*, 35 (November), 464–73.

Herr, Paul M., Peter H. Farquhar, and Russell H. Fazio (1996), "Impact of Dominance and Relatedness on Brand Extensions," *Journal of Consumer Psychology*, 5 (2), 135–59.

Holden, Stephen J.S. and Richard J. Lutz (1992), "Ask Not What the Brand Can Evoke; Ask What Can Evoke the Brand?" in *Advances in Consumer Research*, Vol. 19, John F. Sherry Jr.

and Brian Sternthal, eds. Provo, UT: Association for Consumer Research, 101–107.

John, Deborah Roedder, Barbara Loken, and Christopher Joiner (1998), "The Negative Impact of Extensions: Can Flagship Products Be Diluted?" *Journal of Marketing*, 62 (January), 19–32.

Johnson, M.D. (1986), "Consumer Similarity Judgements: A Test of the Contrast Model," *Psychology and Marketing*, 3 (1), 47–60.

Kahana, Michael J. (2002), "Associative Symmetry and Memory Theory," *Memory & Cognition*, 30 (September), 823–41.

Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (January), 1–22.

Laforet, Sylvie and John Saunders (1999), "Managing Brand Portfolios: Why Leaders Do What They Do?" *Journal of Advertising Research*, 39 (January–February), 51–66.

Loken, Barbara and James Ward (1990), "Alternative Approaches to Understanding the Determinants of Typicality," *Journal of Consumer Research*, 17 (September), 111–26.

Morrin, Maureen (1999), "The Impact of Brand Extensions on Parent Brand Memory Structures and Retrieval Processes," *Journal of Marketing Research*, 36 (November), 517–25.

Nedungadi, Prakash (1990), "Recall and Consumer Consideration Sets: Influencing Choice Without Altering Brand Evaluations," *Journal of Consumer Research*, 17 (December), 263–76.

Olson, James M. and Mark P. Zanna (1993), "Attitudes and Attitude Change," *Annual Review of Psychology*, 44, 117–54.

Posavac, Steven S., David M. Sanbonmatsu, Maria Cronley, and Frank R. Kardes (2001), "The Effects of Strengthening Category-Brand Associations on Consideration Set Composition and Purchase Intent in Memory-Based Choice," in *Advances in Consumer Research*, Vol. 28, Mary C. Gilly and Joan Meyers-Levy, eds. Provo, UT: Association for Consumer Research, 186–89.

Rajagopal and Romulo Sanchez (2004), "Conceptual Analysis of Brand Architecture and Relationships Within Product Categories," *Journal of Brand Management*, 11 (February), 232–47.

Ratneshwar, S., Cornelia Pechmann, and Allan D. Shocker (1996), "Goal-Derived Categories and the Antecedents of Across-Category Consideration," *Journal of Consumer Research*, 23 (December), 240–51.

Roehm, Michelle and Alice M. Tybout (2006), "When Will a Brand Scandal Spillover, and How Should Competitors Respond?" *Journal of Marketing Research*, 43 (August), 366–73.

Schmitt, Bernd H., Nader T. Tavassoli, and Robert T. Millard (1993), "Memory for Print Ads: Understanding Relations Among Brand Name, Copy, and Picture," *Journal of Consumer Psychology*, 2 (1), 55–81.

Scott, Carol A. and Alice M. Tybout (1981), "Theoretical Perspectives on the Impact of Negative Information: Does Valence Matter?" in *Advances in Consumer Research*, Vol. 8, Kent B. Monroe, ed. Ann Arbor, MI: Association for Consumer Research, 408–409.

Skowronski, John J. and Donal E. Carlston (1989), "Social Judgment and Social Memory: The Role of Cue Diagnosticity in Negativity, Positivity, and Extremity Biases," *Journal of Personality and Social Psychology*, 52 (4), 689–99.

Sujan, Mita (1985), "Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgments," *Journal of Consumer Research*, 12 (June), 31–46.

Tversky, Amos (1977), "Features of Similarity," *Psychological Review*, 84 (4), 327–52.

——— and Itamar Gati (1978), "Studies of Similarity," in *Cognition and Categorization*, E. Rosch and B.B. Lloyd, eds. Hillsdale, NJ: Lawrence Erlbaum Associates, 79–98.

Ulhaque, Ehsan and Kenneth D. Bahn (1992), "A Spreading Activation Model of Consumers' Asymmetric Similarity Judgment," in *Advances in Consumer Research*, Vol. 19, John F. Sherry Jr. and Brian Sternthal, eds. Provo, UT: Association for Consumer Research, 782–87.

Weiner, Bernard (2000), "Attributional Thoughts About Consumer Behavior," *Journal of Consumer Research*, 27 (December), 382–87.

Cenk Koçaş & Jonathan D. Bohlmann

Segmented Switchers and Retailer Pricing Strategies

Empirical studies reveal a surprisingly wide variety of pricing strategies among retailers, even among Internet sellers of undifferentiated homogeneous goods, such as books and music CDs. Several empirical findings remain puzzling; for example, within the same market, some small retailers decide to discount deeply, whereas others forgo the price-sensitive switchers and price high. The authors present theoretical and empirical analyses that address these varied pricing strategies. A model of three asymmetric firms shows that under multiple switcher segments, in which different switchers compare prices at different retailers, firm-specific loyalty is not sufficient to explain the variety of pricing strategies. The authors demonstrate that a retailer's strategy to discount deeply or frequently is driven by the ratio of the size of switcher segments for which the retailer competes to its loyal segment size. The relative switcher-to-loyal ratios among retailers explain situations in which a small retailer finds it optimal to price high, despite having few loyals, or to discount and go for the switchers. The results of two empirical studies confirm the model's predictions for varied pricing strategies in the context of Internet booksellers. The analyses also present several implications. A small retailer can sometimes benefit from strategically limiting its access to switchers to soften price competition. A midsized retailer can benefit from targeting its switcher acquisition activities toward its larger rival, given the shallower discounts involved. When most switchers widely compare prices, a large retailer should offer few shallow discounts because other firms will more aggressively discount. The importance of switcher segmentation suggests that managers should carefully measure switching behavior in devising pricing strategies.

Keywords: Internet retailing, pricing strategy, game theory, retailer loyalty, switcher segments

Internet retailing offers consumers considerable choices in terms of where to shop and purchase. For example, a search of several products on mySimon, a popular price comparison engine, reveals more than three dozen retailers for online books, more than 70 retailers offering printers, and more than 100 digital camera retailers. Many of these retailers are small compared with the big players in the category. For example, the Internet book retailer a1Books has less than 1% of the reach (number of viewers who visit the site) of Barnesandnoble.com (B&N). There are dozens of similarly small Internet booksellers that compete with one another and the considerably larger firms of Amazon.com and B&N for a wide assortment of books. What characterizes the price promotion strategies among these many large and small retailers?

In general, frequent discounting by small firms is consistent with extant theory (e.g., Narasimhan 1988; Raju, Srinivasan, and Lal 1990). Customers are conceptualized to possess different information for comparison shopping. Customers who know the price of only one retailer are uninformed or "loyal" customers, whereas customers who know the prices of all retailers are "switchers." Urbany, Dickson, and Kalapurakal (1996) report various reasons more than half of grocery customers compare prices across retailers, and "cherry-picking" customers tend to benefit from going store to store for specials (Fox and Hoch 2005). Because a small firm has relatively few loyal customers, it typically has more incentive than a large firm to discount and attract the switchers. However, some small retailers decide to price high and sell only to their loyal or niche segment of customers. Although this strategy may be appealing for specialized or highly differentiated goods, it is unclear why a small retailer would price higher than a larger rival for a homogeneous good, such as a book. Clay, Krishnan, and Wolff (2001) report these types of "puzzling" Internet retailers in their empirical analyses—namely, small and undifferentiated firms that, nevertheless, charge relatively high prices. In general, Internet retailers exhibit considerable price dispersion as firms discount or price high, often in ways that are inconsistent with theoretical predictions (Pan, Ratchford, and Shankar 2004).

Especially problematic for conventional price promotion theories is the observation of some small retailers with heavy discounting and others with a high-price strategy in the same homogeneous goods market. Previewing our empirical analyses of Internet booksellers, both Worthy and a1Books are small retailers with similar Internet reach and site popularity. They sell many of the same books (a1Books carries 90% of the books carried by Worthy in our sample) and also compete with the larger retailers, such as Amazon.

Cenk Koçaş is Associate Professor of Marketing, Faculty of Management, Sabancı University, Istanbul (e-mail: kocas@sabanciuniv.edu). Jonathan D. Bohlmann is Associate Professor of Marketing, College of Management, North Carolina State University (e-mail: Jon_Bohlmann@ncsu.edu). The authors thank the anonymous *JM* reviewers, Serdar Sayman, Ruth Bolton, Peter Fader, Xing Pan, Dina Mayzlin, and seminar participants at Michigan State University for their helpful comments. Tunga Kiyak provided excellent research assistance. The support of Roger Calantone, Eli Broad University Professor of Business at Michigan State University, is gratefully acknowledged. The authors thank comScore Inc. for supplying the data.

© 2008, American Marketing Association
ISSN: 0022-2429 (print), 1547-7185 (electronic)

Journal of Marketing
Vol. 72 (May 2008), 124–142

com, on a wide variety of books. Despite such similarities, a1Books discounts heavily whereas Worthy typically has high prices that exceed those of both a1Books and Amazon.com. What explains these fundamental pricing differences among small retailers? How frequently do large retailers offer shallow or deep discounts? We present theoretical and empirical analyses that examine various observed price promotion strategies, such as in Internet retailing.

We address two fundamental shortcomings of conventional price promotion models: (1) The duopoly structure can be limited in the various price promotion strategies that are considered, and (2) switchers are usually assumed to be homogeneous in that they know the prices of all retailers. Retailer price promotion strategies from the standard duopoly approach (small versus large retailers) are driven by different loyal segment sizes and homogeneous switchers who compare all prices.[1] We present a model of an asymmetric triopoly (three firms) that includes multiple "switcher segments," in which switchers compare prices at different retailers. Markets involving several asymmetric firms may prove more insightful and reflective of actual market conditions than the conventional duopoly. Furthermore, in the context of multiple retailers, there may be multiple segments of "partially informed" switchers in addition to the typical switchers who compare prices for all retailers. Even in an Internet setting, in which retailers proliferate and many shoppers use price comparison engines, not all switchers will be exhaustive price comparison machines. For example, many online book customers may compare prices between the two larger firms, Amazon.com and B&N, while ignoring smaller retailers such as a1Books. Recent studies indicate that, on average, online shoppers visit only 1.2–1.4 booksellers (Johnson et al. 2004; Montgomery et al. 2004). Even the use of price comparison "shopbots" remains relatively low, given the various cognitive costs of evaluating many alternatives (Montgomery et al. 2004). Consumers are more likely to visit Internet retailers they trust because of an assortment of factors (e.g., Bart et al. 2005). If some customers prefer to compare prices only among trusted stores whereas other customers compare prices from multiple listed retailers, switcher segments with different types of price comparison behavior will occur. Therefore, the realistic assumption is that switchers are segmented because they compare a subset of retailer prices. By including segmented switchers in an asymmetric triopoly, our model derives a more complete set of promotion strategies that are not predicted under asymmetric duopolies.

Prior theoretical models, based largely on the work of Varian (1980) and Narasimhan (1988), report varied results for promotional activity under different conditions. Raju, Srinivasan, and Lal (1990) conclude that the average discount offered by the stronger brand (greater loyalty) is larger than the average discount offered by the weaker brand (less loyalty), but Rajiv, Dutta, and Dhar (2002) reach a different conclusion. Findings also differ about price promotion frequency. Narasimhan (1988) and Raju, Srinivasan, and Lal (1990) conclude that the promotional frequency of the stronger brand is less than that of the weaker brand.[2] Rajiv, Dutta, and Dhar (2002) conclude that high-service stores (as an analogue to a strong brand) promote more frequently under promotional advertising. Small retailers seem to promote with less intensity, though discount prices may have a higher expected payout to smaller stores (Hoch, Drèze, and Purk 1994; Shankar and Bolton 2004). The high versus low pricing dichotomy does not often fit the rich promotional strategies observed empirically (e.g., Bolton and Shankar 2003). There is a need for research that addresses these price promotion variations, particularly for the pricing strategies of small retailers.

We consider a homogeneous goods market with three firms that are asymmetrical in loyals and switchers, and our results clarify the varied price promotion strategies in equilibrium. For example, we show that a firm can adopt a "partitioned" pricing strategy that combines frequent, shallow discounts to compete with a higher-priced firm and infrequent, deep discounts to compete with a lower-priced firm. Our results also explain a situation in which a small firm is better off pricing high with little discounting, despite possessing few loyals. In general, our model not only encompasses a three-firm extension of prior duopoly models but also explains previously ambiguous situations, such as when a retailer has both few loyals and few switchers.

Our specific research questions include the following: How do multiple, asymmetric firms compete for multiple switcher and loyal segments? How do firms differ in their price dispersion, including the frequency and depth of their promotions, in terms of their loyal and switcher customer bases? and Why do some small retailers price high and others offer deep discounts? Our results demonstrate that a retailer's pricing is driven by the ratio of switchers for which the retailer competes to its loyal segment size. The retailers' relative switcher-to-loyal ratios (SLRs) explain when a large or small firm is more or less inclined to discount deeply or frequently or when a small firm with few loyals is better off pricing high.

We test our model's predictions with two empirical studies, using pricing data for numerous Internet book retailers for more than 1600 books. The two studies differ in how an empirical measure is formed for a retailer's SLR (the key variable of the model). The first study is exploratory, using publicly available descriptive data for the retailers. The second study uses comScore data of Internet browsing and purchase data for 100,000 panelists. Compared with the exploratory study, the panel data study has the advantage of more directly measuring switching and loyal behaviors; the disadvantage is that fewer retailers are represented. The empirical results for both studies confirm the model's predictions for retailer pricing based on the SLR, adding a new perspective to previous studies of dispersed prices (Baye and Morgan 2001; Burdett and Judd 1983; Raju, Srinivasan, and Lal 1990; Salop and Stiglitz 1977).

---

[1] Prior models have been applied to both brand and retailer competition. We model retailers given our focus on homogeneous goods and to be consistent with the context of our empirical setting.

[2] Narasimhan (1988) considers several cases in which the frequency and depth of discount changes according to switchers' brand preferences. For homogeneous goods, we compare our results with Narasimhan's basic results.

We organize the remainder of this article as follows: We present our model, highlighting the strategic intuition in detail and contrasting our results with prior models. We also formulate several hypotheses from the model's results. We then describe our empirical studies and the results. We conclude with managerial implications and suggestions for further research.

Model and Analysis

Our model builds on the work of Narasimhan (1988) and Varian (1980). Whereas Varian considers many symmetric firms and Narasimhan studies two asymmetric firms, we analyze a market with three asymmetric firms and multiple switcher segments. Switcher segments among multiple, asymmetric firms are not components of prior models, which makes our research unique. For conventional retailers, it is easy to imagine that not all customers will be informed of all prices because of the high cost associated with searching prices. In an Internet setting, however, search costs are relatively low. We can reasonably assume that at least some customers are highly informed about online prices, even for low-cost items (Carlton and Chevalier 2001; Koçaş 2002; Smith and Brynjolfsson 2001). However, this assumption does not preclude multiple segments of switchers who do not compare prices among every retailer. Even with price comparison engines, not all retailers are listed, and not all customers use price comparison sites (Iyer and Pazgal 2003; Montgomery et al. 2004). Asymmetric awareness of retailers by price-sensitive switchers may lead to different pricing strategies (Pan, Ratchford, and Shankar 2004).

The fundamental intuition behind our model is that multiple switcher segments can lessen price competition among firms (for a model of price competitiveness in the context of consumer search costs, see Lal and Sarvary 1999). Firms with greater motivation to discount, because of a smaller loyal segment size and/or a greater number of switchers to potentially serve, will more actively compete for the fully informed switchers. This leaves firms with fewer relevant switchers for a given loyalty size to focus on their loyals and the subset of switchers who consider them in their price comparison search. In such cases, prices will typically be higher than they would be if the firm were to compete for all switchers. Other firms discount less in reaction to these higher prices, and thus the severity of price competition becomes less overall. Under some conditions, the firm with the fewest loyals can be the highest-priced firm, a result that asymmetric loyalty by itself cannot implement. Our model predictions reflect a wide variance in retailers' price dispersions, consistent with empirical observations.

Consider a market for a homogeneous good, such as a specific book or music CD, sold by three retailers. On the demand side, each customer purchases a single unit of the good if it is offered at or below the reservation price r, which is assumed to be homogeneous for all customers and across all retailers.[3] We model different segments of both loyal customers who buy from their preferred retailer as long as the price does not exceed r and switchers who comparison shop. The loyal customers are faithful to only one firm; the number of customers who are loyal to Firm i is n_i. Without loss of generality, we assume that $n_1 > n_2 > n_3$. We also assume the existence of two switcher segment sizes, s_{123} and s_{12}, whose members are not loyal to any firm but rather buy from the lowest-priced retailer among those they compare. Switcher segment s_{123} is fully informed, meaning that its members compare prices quoted by all three firms. Switcher segment s_{12} compares prices quoted only by the larger Firms 1 and 2. Although these partially informed switchers do not compare all retailer prices, they still compare the prices from the best-recognized retailers, which may have the largest reach and the most active communication channels.[4] Firms 1 and 2 would be like B&N and Amazon.com (the biggest online booksellers), whereas Firm 3 would be like a1Books. Although some switchers will compare prices at all three retailers (s_{123}), a1Books has less than 1% of B&N's reach, so there are likely to be other switchers who are unfamiliar with a1Books and compare only Amazon.com and B&N (s_{12}).

We normalize the market size to one ($n_1 + n_2 + n_3 + s_{12} + s_{123} = 1$) without loss of generality. Although the segment sizes and the reservation price are common knowledge, because of imperfect addressability and targetability, retailers cannot price discriminately (Blattberg and Deighton 1991; Chen, Narasimhan, and Zhang 2001). All firms face constant fixed and marginal costs, which we assume to be zero without loss of generality (Iyer and Pazgal 2003; Narasimhan 1988; Raju, Srinivasan, and Lal 1990). Overall, the model and its assumptions are similar to Narasimhan's (1988) model and other related models, except for analyses of three firms and two switcher segments.

In an effort to capture switcher segments, firms have an incentive to undercut the price of other firms that are competing for those switchers, a tendency that results in a downward push in prices. Motivation also exists to price at the reservation price, in case the switchers cannot be served with a lower price. A retailer's "minimum price" makes the firm indifferent between selling only to its niche of loyal customers at the reservation price and selling to its switcher segments, given it is the lowest-priced firm at the minimum price. A firm will never discount below its minimum price, because it could then do better by focusing on its loyal customers. A firm's smaller loyal segment size and larger

---

[3] A common reservation price is a typical assumption (Iyer and Pazgal 2003; Narasimhan 1988; Raju, Srinivasan, and Lal 1990; Varian 1980). Narasimhan (1988) considers a case in which consumers have different reservation prices for brands.

[4] Consumers may choose not to compare a specific retailer's prices for several reasons, such as a lack of awareness or a lack of trust in the retailer. Our definition of switcher segments is based fundamentally on the set of retailers among which the switcher price compares. A dynamic extension to our model with endogenous segment sizes may take different forms depending on whether price comparisons are dependent on awareness or trust. We thank a reviewer for raising these distinctions.

switcher segment size provide a greater incentive to forgo selling only to the loyals. We define a firm's ratio of switcher-to-loyal segment sizes, ϕ_i, as $\phi_3 = s_{123}/n_3$, $\phi_2 = (s_{123} + s_{12})/n_2$, and $\phi_1 = (s_{123} + s_{12})/n_1$. Under segments s_{123} and s_{12}, Firms 1 and 2 compete for the same switchers, so Firm 2 always has a higher SLR (or ϕ_i) than Firm 1, given Firm 2's smaller loyal segment. However, Firm 3 can only capture switcher segment s_{123}, so its SLR may be higher or lower than that of both firms. Thus, Firm 3 may discount deeply to capture the switchers or play only its niche loyal segment and charge higher prices. Without multiple switcher segments, the relative order of ϕ_i for these firms would strictly be a function of the loyal segment sizes, as in prior models. Thus, our recognition of segmented switchers changes the nature of a retailer's incentive to discount.

A more general consideration of segmented switchers (adding s_{23} and s_{13} switcher segments) does not change the model's intuition or conclusions. Additional switcher segments modify the degree of a retailer's relative incentive to discount, but the pricing equilibrium remains a function of retailers' SLRs and their rank order. Generalizing to more than three asymmetric retailers also gives results that are similar to a triopoly (details are available on request). Thus, our model focuses on three firms and two switcher segments for relative tractability.

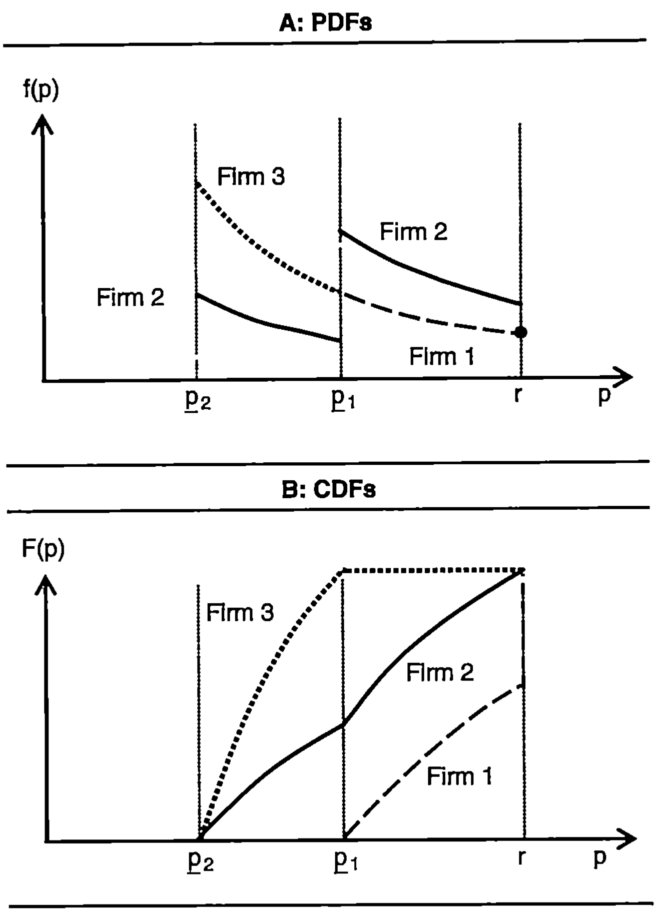
Equilibrium Price Promotion Strategies

The tension of selling to switchers and loyals results in a lack of pure strategies typical of such models. Therefore, we solve for a mixed-strategy equilibrium that depends on the relative sizes of the loyal and switcher segments. A mixed strategy can be interpreted as arising from a retailer's uncertainty about the pricing decisions of competing retailers (Gibbons 1992). Equilibrium prices are defined by a probability density function (PDF) that indicates the range of prices a retailer may charge (the retailer's price support). We find that equilibrium price promotion strategies fundamentally depend on the retailers' relative SLRs. Thus, we present the mixed-strategy equilibrium for the case that is more typical of prior models, in which the large Firm 1 has the lowest ϕ_1, as follows:

> P_1: When $\phi_3 > \phi_1$, Firms 1 and 3 have mutually exclusive price ranges that, when combined, form Firm 2's price range. Firm 2 competes with Firm 3 in the lower part of its price range and with Firm 1 in the upper part of its price range (for a proof, see the Appendix).

The PDF and the cumulative distribution function (CDF) for the equilibrium prices appear in Figure 1. Firm 2 has greater motivation to compete for s_{123} than Firm 1, and because Firm 3 can only capture s_{123}, the competition between Firms 2 and 3 is fairly intense. The low prices Firm 2 quotes while trying to capture s_{123} make it easier for Firm 2 to sell to s_{12}. Thus, there is no guarantee that Firm 1 will capture any switchers, even at its minimum price, because the other two firms already compete for s_{123} at lower prices. Therefore, Firm 1 is less inclined to discount. The more intense competition between Firms 2 and 3 makes the lower bound of Firm 1's price range move up to the point in equilibrium at which Firm 1 no longer directly competes for s_{123}. Thus, Firm 1's and Firm 3's equilibrium price ranges do not overlap. Moreover, when Firm 2 prices below Firm 1 to compete with Firm 3 for segment s_{123}, it also always serves segment s_{12}. That is, in the price region in which only Firms 2 and 3 compete, s_{12} effectively becomes a component of Firm 2's loyal segment. Because an effectively larger loyal segment means that Firm 2 now has more to lose from price cuts, the lowest price that Firm 2 can profitably quote increases, which lessens the severity of the price competition with Firm 3. Thus, the lowest price support for Firms 2 and 3 rises above the minimum prices (p_1^{min}) of both firms.

FIGURE 1
Equilibrium Prices in P_1, Illustrated for $s_{123} = s_{12}$


A: PDFs

B: CDFs

These results emerge mainly because of asymmetry in *both* the switcher and the loyal segment sizes. The lowest prices are higher in our model than in Varian's (1980), Narasimhan's (1988), and other models that lack multiple switcher segments. The existence of s_{12} as a switcher segment that omits Firm 3 leads to higher average prices for Firms 2 and 3, and their competition for segment s_{123} leads to higher average prices for Firm 1. At the extreme of $s_{12} = 0$ (where $\phi_3 > \phi_2$), the results represent a three-firm extension of Narasimhan's model, similar to other models of asymmetric oligopoly (Baye, Kovenock, and De Vries 1992; Koçaş and Kiyak 2006).

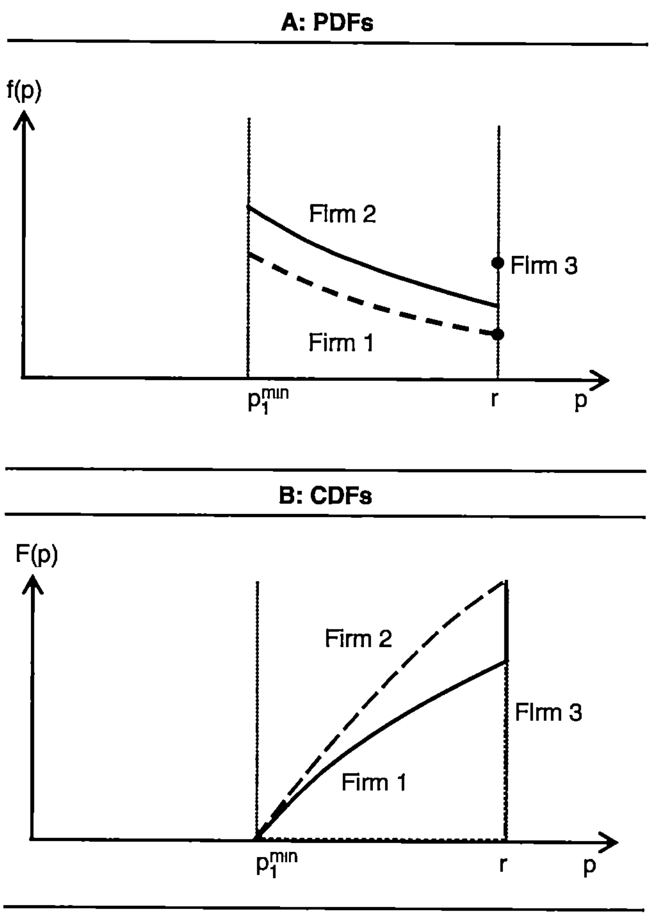
We find that the equilibrium of P_1 does not change, regardless of whether Firm 3's SLR exceeds that of Firm 2. The intuition is straightforward; the lower price support for

Firms 2 and 3 rises above their respective minimum prices. Therefore, the equilibrium is unaffected by whether the minimum prices of Firms 2 and 3 switch order, because neither is part of the equilibrium price ranges for those two firms. Furthermore, although Firm 2 may have the highest SLR, it still sets its lower price support partially in response to Firm 1's higher support, because Firm 1 forgoes competition for s_{123}. Firm 1's main issue is that both Firms 2 and 3 have higher SLRs, regardless of whose is greater.

P_1 contrasts remarkably with the case in which Firm 3 has the (weakly) lowest SLR, or $\phi_3 \leq \phi_1$ (recall that $\phi_2 > \phi_1$). In this situation, Firm 3 lacks sufficient discount motivation to capture its switcher segment and thus becomes a high-priced niche retailer focusing on its loyal segment. Because Firm 1's SLR is not the lowest, it finds that it is worthwhile to play for s_{123}. Segments s_{12} and s_{123} implicitly combine into a single switcher segment that only Firms 1 and 2 serve. The resultant equilibrium is defined as follows:

> P_2: When $\phi_3 \leq \phi_1$, only Firms 1 and 2 discount, whereas Firm 3's price is set at the reservation price r (for a proof, see the Appendix).

Figure 2
Equilibrium Prices in P_2



A: PDFs

B: CDFs

The PDFs and CDFs for the equilibrium prices appear in Figure 2. The lower price support is defined by Firm 1's minimum price because Firm 2 never needs to discount below this to compete for the switchers. Firm 3 is content to be a niche player, pricing high to serve only its relatively small loyal segment. Given that $n_1 > n_3$, the P_2 condition that $\phi_3 \leq \phi_1$ essentially represents the case in which s_{123} is small compared with s_{12} ($s_{123} = 0$ will always meet the condition of P_2). Firm 3 recognizes that most of the switchers compare prices only between Firms 1 and 2, so Firm 3 becomes a high-priced retailer.

Model Discussion

Before turning to the empirical validation of our model, we describe in more detail the firms' profits and prices in the presence of segmented switchers (for a summary, see Table 1). With a partially informed switcher segment s_{12} included, average prices and profits are generally higher because the price supports increase as Firm 1 "abandons" s_{123} under P_1. Thus, multiple switcher segments tend to reduce competition because not all firms are able to capture all switchers.

The relative sizes of the switcher segments shape the nature of competition between the firms. Under P_1, as Figure 3, Panel A, depicts, Firm 1's discount deepens as s_{12} exceeds s_{123} (when the total number of switchers is held constant). However, when the size of the fully informed segment increases, so does the severity of competition between Firms 2 and 3. This increased competition for segment s_{123} pushes Firm 2's and Firm 3's discounts deeper (Figure 3, Panel B). At the same time, it forces Firm 1 to offer shallower discounts because Firm 2 already serves more of segment s_{12} (in expectation) with its deeper discounts. Thus, the composition of switchers fundamentally affects when retailers should offer frequent or infrequent deep or shallow discounts.

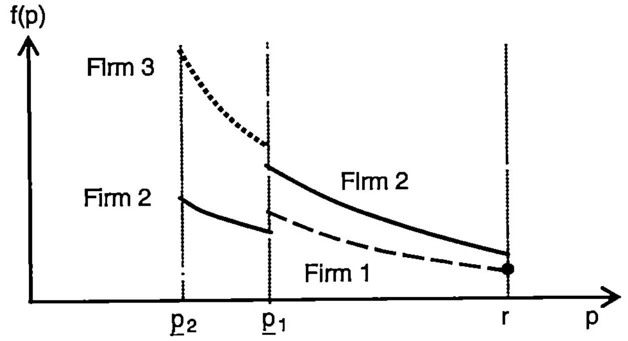
When we compare firms' profits as the composition of switcher segments varies, we first observe that Firm 1 is indifferent because it always receives the guaranteed profit (n_1r). Firm 2's profits increase when s_{12} increases at the expense of s_{123}, given that segment s_{12} acts as a second-degree loyal segment to Firm 2. A more notable result pertains to the profits of Firm 3 under P_1. Firm 3 can benefit

TABLE 1
Summary of Equilibrium Results in P_1 and P_2

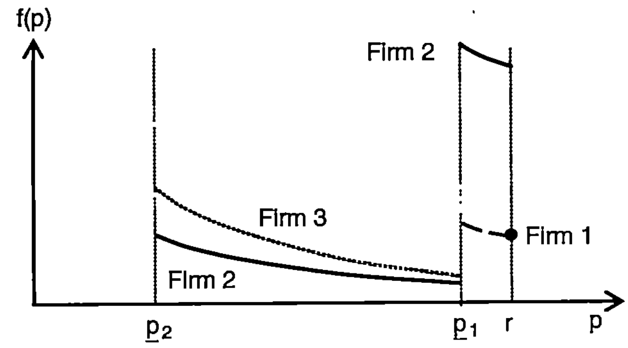
	P_1 ($\phi_3 > \phi_1$)		P_2 ($\phi_3 \leq \phi_1$)	
	Profits	**Lowest Price**	**Profits**	**Lowest Price**
Firm 1	n_1r	$\underline{p}_1 > p_1^{min}$	n_1r	p_1^{min}
Firm 2	$(n_2 + s_{12})\underline{p}_1 > n_2r$	$\underline{p}_2 \geq p_2^{min}$	$(n_2 + s_{12} + s_{123})p_1^{min} > n_2r$	p_1^{min}
Firm 3	$(n_3 + s_{12} + s_{123})\underline{p}_2 > n_3r$	$\underline{p}_2 \geq p_3^{min}$	n_3r	r

FIGURE 3
Depth and Frequency of Promotions

A: Size of the Fully Informed Switcher Segment Is Smaller Than the Size of the Partially Informed Switcher Segment ($s_{123} < s_{12}$; P_1)



B: Size of the Fully Informed Switcher Segment Is Larger Than the Size of the Partially Informed Switcher Segment ($s_{123} > s_{12}$; P_1)



from a positive s_{12} customer base as long as the higher price support more than compensates for the loss of some s_{123} customers as s_{12} increases. In general, Firm 3's profits do not reach a maximum when all switchers belong to s_{123}.[5] Thus, Firm 3 may strategically choose to limit its exposure to switchers (lowering s_{123} and increasing s_{12}) to make Firm 2 focus more on its competition with Firm 1. That is, Firm 3 benefits from a more balanced distribution of switchers, a condition that places Firm 2 in a more balanced consideration for both switcher segments. Thus, our model complements Iyer and Pazgal's (2003) and Baye and Morgan's (2001) results, which pertain to retailers belonging (or not) to Internet agents that facilitate price comparisons. By considering various-sized switcher segments, our model points to a small retailer potentially benefiting from some degree of price comparison strictly between the larger or better-known firms. This result further highlights the competitive nature of the asymmetric triopoly. Firm 3's prices and profits depend critically on how Firms 1 and 2 compete with each other for its switcher segment.

When the relative size of the fully informed switcher segment is small, Firm 3 may prefer to ignore the switchers completely and offer no discounts at all (P_2). This result represents an important distinction between our segmented-switcher model and other models of asymmetric loyalty. When multiple switcher segments exist, loyal segment sizes are not sufficient to describe the nature of competition among the various firms. The relative sizes of the switcher segments must also be considered because they affect the SLRs (ϕ_i) among the firms. Thus, a small firm may discount or price high to play its niche, depending on its relative SLR.

Comparison with Prior Models

Our model complements various prior findings for promotion frequency and depth. In terms of promotion frequency, in general, P_1 results follow the work of Narasimhan (1988) and Raju, Srinivasan, and Lal (1990) in that they predict that a strong retailer (one with greater loyalty) promotes less frequently. However, our model also predicts that a strong retailer can promote more frequently if it has relatively more interest in the switchers than the weaker retailers have (e.g., Firm 1 promotes more often than Firm 3 under P_2). Narasimhan notes that a strong brand may be promoted more frequently (with shallower discounts) if switchers prefer that brand, whereas our results pertain to the switcher segment sizes rather than brand or retailer preference. Rajiv, Dutta, and Dhar (2002) conclude that a high-service (i.e., strong) store offers advertised sales more frequently, albeit to build traffic. The results of P_2 concur with the predictions of Rajiv, Dutta, and Dhar, even in the absence of traffic considerations, suggesting that a relatively larger size of switchers targeted by high-service stores can explain a higher frequency of advertised sales.

In the case of promotion depth, recall that Narasimhan (1988), Raju, Srinivasan, and Lal (1990), and Rajiv, Dutta, and Dhar (2002) all examine asymmetric duopolies. Under their respective models, Narasimhan suggests that the two stores offer the same depth of discounts (for indifferent switchers); Raju, Srinivasan, and Lal predict that the stronger (or larger) store should offer deeper discounts; and Rajiv, Dutta, and Dhar predict that the stronger store should offer shallower discounts. Our model of segmented switchers predicts the conditions under which a store may offer deeper or shallower discounts. More specifically, a strong store may indeed offer shallower discounts if it has the least motivation to promote, as characterized by its smaller SLR (P_1). A strong store may also offer the same discount depth as a smaller store if no other weak firms compete for switchers (P_2; Firms 1 and 2). A strong store may even offer deeper discounts than a weak store if its share of switchers is relatively larger than that of the weak store (P_2; Firms 1 and 2 compared with Firm 3). Thus, our three-firm model with segmented switchers encompasses various cases of prior duopoly models.

Iyer and Pazgal (2003) present a model of many firms (as in Varian 1980) with switchers and asymmetric loyalty,

---

[5] Firm 3's profits for at least some positive region of s_{12} are higher than its profits at $s_{12} = 0$. More specifically, as long as the total number of switchers remains above a modest threshold, Firm 3's profit has a positive slope at $s_{12} = 0^+$, such that there is at least some positive s_{12} in which Firm 3 is better off.

but they include, at most, two types of firms. Iyer and Pazgal also consider "partial loyals," who behave as loyals with a lower reservation price. In contrast, we model three asymmetric firms with multiple switcher segments. Therefore, some of our results differ from those of Iyer and Pazgal and other models that lack segmented switchers. In particular, we predict the existence of a high-priced niche strategy for a small retailer, which we empirically observe. Other models do not make such predictions for homogeneous goods.

Our model also offers some unique insights into price strategies when a retailer competes on multiple fronts. The novel discounting behavior of Firm 2 in P_1 completely draws from the three-firm asymmetry with segmented switchers. Firm 2 offers two types of discounts: deep discounts to compete with Firm 3 and shallow discounts to compete with Firm 1. Furthermore, as Figure 3 depicts, depending on the relative sizes of the switcher segments, the frequency of deep and shallow discounts varies. Competing with different firms for different switcher segments enables a retailer to develop a partitioned pricing strategy, in which the depth and frequency of discounts vary on each competitive front.

Model Predictions

To demonstrate that our model can explain pricing behavior in a rich empirical context, we present testable model predictions and examine them with pricing data from Internet book retailers. Our model yields price distributions as equilibrium strategies that depend on the relative SLRs of the firms.[6] Although summary statistics that describe the promotion strategies can be used to test the fit of pricing data in general, a more robust method is to compare the empirical and theoretical price distributions themselves. We offer predictions based on summary statistics in H_1 and predictions based on the price distributions in H_2 and H_3.

> H_1: On average, retailers with a higher SLR have (a) lower average prices, (b) higher standard deviations in prices, (c) more frequent discounts, (d) a higher maximum discount frequency, (e) greater discount depth, and (f) a greater maximum discount depth.

Under the model's propositions, firms with a higher SLR will have lower average prices (H_{1a}) and a higher standard deviation in prices (H_{1b}) in relation to price deviation from the reservation (regular) price. The frequency of price discounts is related to the likelihood that a retailer will price below the maximum retail price. Under P_1, Firms 2 and 3 will discount more frequently than Firm 1, and under P_2, a firm with a higher SLR will promote more frequently than a firm with a lower SLR. Overall, a firm with a higher SLR will discount at least as often as a firm with a lower SLR and therefore will discount more frequently (H_{1c}) and have a higher maximum discount frequency (H_{1d}), on average.

From our model, a firm with a lower price support will tend to have a greater depth of discount. Under P_1, for example, Firm 2 has a greater probability of discounting at a lower price than Firm 1, as evidenced in the CDF plot of Figure 1, Panel B. Across all possible conditions of the model's propositions, we expect to observe that, on average, retailers with higher SLRs will have greater discount depth (H_{1e}) and a greater maximum discount depth (H_{1f}).

Although summary statistics of retailers' prices can indicate whether the SLR explains price promotions, a more rigorous test would consider the entire price dispersion curve. Theoretically, we have clear predictions about how equilibrium price distributions should appear (Figures 1 and 2). We test whether the empirical price distributions vary as predicted by analyzing stochastic dominance. For any two CDFs, $F_j(x)$ first-order stochastically dominates $F_i(x)$ iff $F_j(x) \leq F_i(x)$, $\forall x$. In other words, if Firm j's price CDF lies nowhere above and somewhere below Firm i's distribution, Firm j first-order dominates. In Figure 1, Panel B, for example, Firm 1 first-order stochastically dominates Firms 2 and 3, and Firm 2 dominates Firm 3. First-order dominance is a fairly strict standard when considering empirical price distributions that may include shocks or random error. Second-order stochastic dominance is similar except that it considers the deficit functions, or the integral of the CDF. Second-order stochastic dominance holds under first-order dominance, but not vice versa. Our model predicts that, on average, a firm with a lower SLR will stochastically dominate a firm with a higher SLR.

> H_2: The price distributions of retailers with a lower SLR first- and second-order stochastically dominate firms with a higher SLR.

H_2 addresses the relationship between SLR and price distributions among all retailers, and H_3 more narrowly focuses on retailers with the smallest loyal segment size (Firm 3). Small retailers with a high SLR should discount heavily (P_1), whereas small firms with a low SLR should price high (P_2). Therefore, we can distinguish the pricing strategies of small retailers according to their SLR values.

> H_3: The price distributions of small retailers with a high SLR are first- and second-order stochastically dominated by small firms with a lower SLR.

Empirical Methodology Overview

We test the predictions of our modeling effort by using pricing data collected for online book retailers. A book is a homogeneous product that is uniquely identified by an ISBN (International Standard Book Number), a classification that is widely used and recognized by both customers and sellers. To ensure that the online retailers in our data set

---

[6] Mixed strategies in pricing are usually assumed to be observable through temporal price dispersion (Narasimhan 1988; Varian 1980) that is pooled across multiple products and retailers (e.g., Iyer and Pazgal 2003; Raju, Srinivasan, and Lal 1990). Data for multiple products (in our case, books) represent repeated observations of a mixed pricing strategy over time. Our analysis is consistent with this view in that we define "discount" in terms of temporal price changes, and we consider a retailer's average discounting behavior across books. Therefore, the hypotheses state the overall average effects we should observe given both propositions taken in tandem. On a case-by-case comparison of retailer pairs, all the hypotheses strictly hold, with a single technical exception to H_{1b}, H_{1e}, and H_2 (Firms 2 and 3 when Firm 2 has the higher SLR under P_1). In its most general form, a mixed strategy is a probability distribution over pure strategies, such that price dispersion may be reflected across books as well as time (for a discussion of mixed-strategy interpretations, see Gibbons 1992).

share at least a common switcher segment, we collected daily data on book prices from all the retailers listed on the price comparison site mySimon. In the period during which our data were collected, mySimon was the leading price comparison engine, with 80% of the price comparison site visit market share (Media Metrix and Nielsen//NetRatings; see also Allen and Wu 2002; Koçaş 2002). Our consideration of booksellers listed on mySimon establishes the existence of a "fully informed" switcher segment (s_{123} in our model), without precluding the existence of other switcher segments. Prices reported on mySimon are taken directly from each retailer's Web site, reflecting the book's offered price to all shoppers at that time, net of shipping and handling costs.

Internet Retailer Pricing Data

We compiled a sample of 2207 books from various sources.[7] From June 2001 to August 2002, we collected daily price data on these books from all the bookstores returned by a search on mySimon. Eliminating duplicates and books not carried by any of the mySimon bookstores left 1640 books. We prefer that price changes in the data reflect the mixed-strategy discounting of retailers rather than systematic changes from price shocks (e.g., net price changes due to a switch over to free shipping) or reservation price changes. Price data collected over longer periods have a greater risk of being contaminated by systematic price adjustments. To alleviate this problem, our empirical analyses use daily pricing data on the 1640 books for a period of 26 days in June 2002. This month represents the most recent period for which we have complete data from all the retailers and no systematic price adjustments. This 26-day window also spans a relatively short period to minimize reservation price adjustments.[8]

In our June 2002 sample period, 28 unique mySimon retailers carried at least 1 of the 1,640 books. All books in the sample were offered by multiple retailers. To compare retailers that carry a similar assortment of books, we ranked the booksellers with respect to the percentage of books they carried from the total list of 1640. We observed a gap between the top 14 retailers, which carried at least 40% of the books, and the bottom 14 retailers, which carried, at most, only 9% of the books. On the basis of this observation, we limit our analyses to the 14 retailers that carried at least 40% of the books. The final sample of retailer prices has 392,245 total observations, or an average of approximately 28,000 per retailer.

Price Promotion Measures and Dispersion

For each retailer, we calculate various measures of price-discounting activity. We do so with the view toward price discounts reflecting price changes over time. During June 2002, each book had a sequence of daily retailer prices. Average price for any retailer is the average of all normalized prices for that retailer; we calculate normalized prices by dividing any given book price by the highest price quoted for that book by any one of the 14 retailers in the sample period. The standard deviation of the normalized prices in each sequence, averaged over all the books carried by the retailer, gives the average standard deviation. For each retailer, we also count the number of price changes in our daily data. We then average that figure across all the books the retailer sold to calculate the average number of price changes (discount frequency). We find the maximum by taking the highest number of price changes for the retailer across books. We calculate the retailer's absolute changes in normalized prices, averaged across all books, to obtain the average and maximum depth of discounts across books for that retailer.

The summary retailer price statistics appear in Table 2. Figure 4 presents the empirical CDF of normalized prices for the 14 retailers. The numbers in parentheses next to each retailer in the figure refer to the same ordering as in Table 2.[9]

Study Procedure

Our objective is to observe how the relative SLRs of the retailers affect their pricing strategies. To form an empirical proxy of the SLR, we need indicators of retailer-specific loyal and switcher segment sizes. We conduct two studies. The first is an exploratory study using readily available descriptive statistics for the retailers. The advantage of this approach is that data for all 14 retailers can be obtained, and the disadvantage is that aggregate retailer descriptives may inadequately proxy the SLR. The second study uses a large Internet panel database that tracks all Internet site visits and purchases for 100,000 panel members. This study's advantage is the rich set of microlevel data to form switcher and loyalty measures. However, the database contains information on a more limited set of Internet booksellers. Given the advantages and disadvantages of each study, we present both as independent tests of our model predictions.

To test H_1, we compare the pricing statistics with retailer SLR values. H_2 and H_3 relate our model predictions

---

[7] These sources consisted of four book lists that were publicly available on the Internet: (1) 140 books from the *Publisher's Weekly* bestsellers list on June 4, 2001; (2) 730 books from "One Book List: Collaborative Books to Read," a list compiled by the Usenet newsgroup (see rec.arts.books); (3) 60 books from the Latest Acquisitions by Government Environmental Library 2001; and (4) 1277 books bought by Sidney Sussex University Library in 2001.

[8] Empirical studies strive to use a "large enough" number of periods to capture temporal realizations of the mixed strategy. Our number of periods is similar to that used by others (e.g., Clay, Krishnan, and Wolff 2001; Raju, Srinivasan, and Lal 1990), though we use daily rather than weekly data to minimize systematic price adjustments. We examined three- and six-month versions of the data set (87 and 184 days, respectively), and the overall results and conclusions were the same. We then repeated our analyses after eliminating bestseller books from the sample. Bestsellers typically have more price dispersion and are more likely to be traffic generators than nonbestsellers (Clay, Krishnan, and Wolff 2001). Our hypotheses remain significantly supported if bestsellers are eliminated.

[9] Figure 4 shows normalized prices for all books across all days. From our model, a retailer that stochastically dominates another retailer for one book should do so for all books (only the reservation price varies across books). Thus, we examine price distributions across both books and time to test for stochastic dominance.

TABLE 2
Summary Statistics for the Retailer Prices

Retailer	Average	SD	Average Frequency	Maximum Frequency	Average Depth	Maximum Depth
Double Discount	.669	.0097	2.150	5	.094	.530
TextatCost	.749	.0017	.048	3	.003	.223
eCampus	.744	.0115	.643	17	.034	.718
a1Books	.749	.0008	.124	1	.009	.363
Books-A-Million	.834	.0017	.060	5	.003	.251
1Bookstreet.com	.935	.0027	.048	2	.002	.100
Amazon.com	.829	.0045	.073	5	.004	.260
B&N	.870	.0137	.206	4	.012	.351
Page One	.742	.0000	.120	2	.000	.045
BookVariety.com	.869	.0000	.000	0	.000	.000
Words Worth Books	.876	.0046	.140	2	.006	.109
Varsity	.919	.0014	.034	1	.002	.091
Worthy	.898	.0000	.000	0	.000	.000
Sam Goody	.773	.0007	.020	1	.001	.066
Retailer average	.818	.0038	.262	3	.012	.222

to measures of stochastic dominance. Figure 4 shows that at the lower price levels, most firms have a small percentage of observed prices. This observation makes a strict test of stochastic dominance problematic because a retailer that may have highly discounted a few books on a few days (of 28,000 prices quoted by the average retailer) distorts the notion that the same retailer may have relatively high prices for all other books across all days. Furthermore, the retailers' CDFs may cross in closely spaced price regions, meaning that neither firm stochastically dominates the other. Therefore, to test H_2 and H_3, we eliminate the 5% lowest prices quoted by each retailer and consider retailer pairs in which stochastic dominance is evident.[10]

Exploratory Study with Retailer Descriptive Data

The model's findings relate retailer pricing to the SLRs (H_1 and H_2) and reveal different strategies for small retailers depending on their SLR values (H_3). To identify the relatively small retailers, we examine reach and link popularity, which capture a retailer's number of visitors and Internet connectivity. Reach measures the number of users who visit a given Internet site and can be obtained from Alexa. Link popularity refers to the total number of links that a site has across the Internet. We measure link popularity as the sum of all retailer links found by the search engines AlltheWeb, AltaVista, Google/AOL, HotBot/Inktomi, and MSN. More specific to booksellers, we measure book-link popularity as the number of sites that link to a book retailer and contain the word "book" on the same page in which the link appears, which we determined using AltaVista. Table 3 presents the reach and popularity measures for the retailers. As we expected, Amazon.com and B&N are the largest retailers by these measures. Books-A-Million has the third-largest reach and popularity, with a book-link popularity comparable to B&N. The other 11 retailers are all relatively small in comparison.

SLR Proxy

A proxy for the SLR should capture a retailer's relative proportion of switchers and loyal customers. Recent studies have pointed to page views per user as a measure of online loyalty (Demers and Lev 2001; Goldfarb 2002). Page views per user, which we obtained from Alexa, are the average number of unique pages viewed per user by site visitors each day. Research has examined the drivers of trust and loyalty in online shopping, and Web site navigation issues arise as important factors (e.g., Bart et al. 2005; Shankar, Smith, and Rangaswamy 2003; Srinivasan, Anderson, and Ponnavolu 2002). Navigating Web site information for in-depth content increases satisfaction and loyalty (Shankar, Smith, and Rangaswamy 2003). Demers and Lev (2001) and Danaher, Mullarkey, and Essegaier (2006) find that page views per user are positively correlated with the frequency of visits, and frequent visitors tend to be more likely to purchase (Moe and Fader 2004a). More page views can also be associated with "lock-in," or an unwillingness to switch sites (Bucklin and Sismeiro 2003).[11] Thus, the page-views-per-user metric represents a reasonable proxy for loyalty, given its association with in-depth navigation and

---

[10] The lowest 5% of a retailer's prices represent prices that are at least eight standard deviations (in intertemporal terms) below the average price for each retailer. This procedure reflects the concept of "almost" stochastic dominance (see Leshno and Levy 2002).

[11] Several studies (e.g., Bucklin and Sismeiro 2003; Danaher, Mullarkey, and Essegaier 2006; Johnson, Bellman, and Lohse 2003; Johnson et al. 2004) note a learning effect whereby frequent visitors become more efficient in browsing over time. However, this does not mean that frequent (i.e., more loyal) visitors browse less than infrequent visitors. Danaher, Mullarkey, and Essegaier (2006) show that more frequent visitors view more pages.

FIGURE 4
Retailer CDFs for All Books

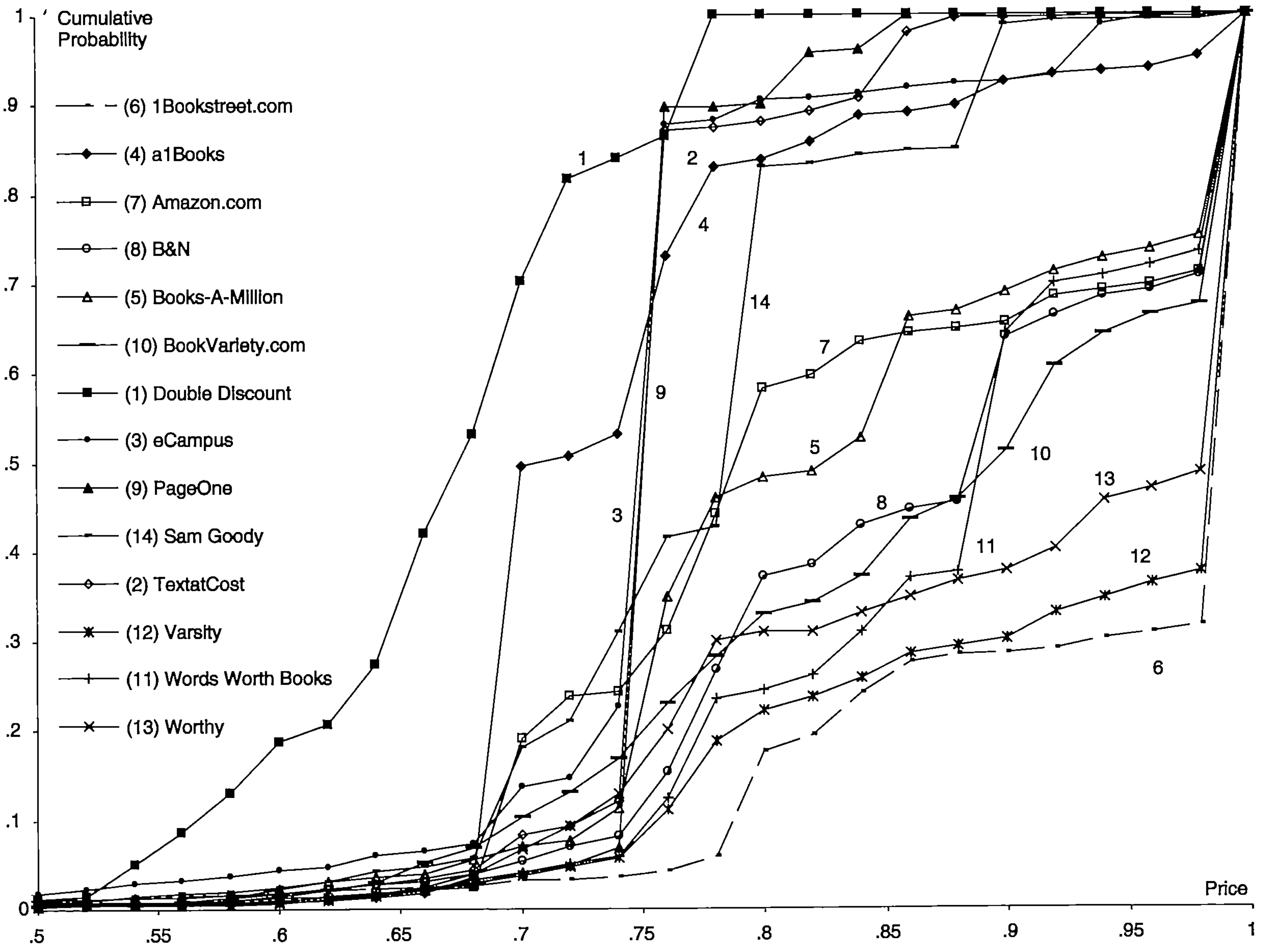


TABLE 3
Exploratory Study Measures for the Retailers

Retailer	Reach/ Millions per Day	Link Popularity	Book-Link Popularity	Page Views per User	Search Share	Normalized Page Views	SLR	SLR Tier
Double Discount	16.5	16,371	100	2.6	1.0	.377	2.65	High
TextatCost	24.5	61,585	2897	3.3	1.0	.478	2.09	
eCampus	75.5	37,418	1998	3.4	1.0	.493	2.03	
a1Books	6.3	25,693	4173	3.5	.7	.507	1.38	
Books-A-Million	81.0	164,506	56,890	5.3	.9	.768	1.17	Medium
1Bookstreet.com	6.65	34,239	6722	2.7	.4	.391	1.02	
Amazon.com	23,600.0	57,347,588	3,676,495	6.8	1.0	.986	1.01	
B&N	1775.0	839,878	46,155	6.2	.9	.899	1.00	
Page One	1.5	9925	274	2.4	.3	.348	.86	
BookVariety.com	1.0	184	24	1.0	.1	.145	.69	Low
Words Worth Books	2.85	20,409	1309	2.6	.2	.377	.53	
Varsity	8.85	31,589	1923	3.9	.3	.565	.53	
Worthy	4.4	19,999	206	1.8	.1	.261	.38	
Sam Goody	63.0	38,012	271	6.9	.1	1.000	.10	
Retailer average	n.m.	n.m.	n.m.	3.7	.57	.54	1.10	

Notes: n.m. = not meaningful, given extraordinary values for Amazon.com and B&N.

loyalty behaviors typically absent from price-sensitive switchers.

To form an indicator of the relative switcher generation potential of the retailers, we observe that a bookstore that wants to attract more switchers may be listed more often in a price comparison query. For a particular set of price comparison sites, we formulate a retailer's search share as the proportion of price comparison sites that return prices for that retailer. Greater search share means that the retailer is listed on more book price comparison sites and thus may attract more switchers from various segments. Although search share does not differentiate switcher segment sizes, it gives a firm-specific indication of switcher intensity (SLR incorporates the sum of a retailer's switcher segment sizes). We use the nine top-ranked book price comparison sites and mySimon to formulate a list of price comparison sites for books.[12] There are several reasons a firm might be listed in a price comparison engine. Sometimes, a price comparison site uses a retailer's prices without specific action on the part of the retailer. In addition, relatively high-cost retailers are not necessarily absent from price comparison sites. A firm with comparably higher prices may want greater access to all potential buyers, and therefore it may share its pricing data with the search engine (Iyer and Pazgal 2003). Whether retailers are high priced or low priced, those listed in price comparison sites have access to a greater number of switchers. Therefore, the degree of participation in price comparison engines serves as an indicator of the retailer's access to switchers who become informed of its prices.

Given these measures, we form a proxy measure of SLR (ϕ_l) as the ratio of search share (a switcher metric) to the normalized page views per user (a loyalty metric). Values appear in Table 3, which orders the retailers by decreasing SLR. Note that the three largest retailers—Amazon.com, B&N, and Books-A-Million—have SLRs in the middle range. For the smaller firms, the SLR separates those with a high SLR (e.g., eCampus) from those with a low SLR (e.g., Varsity), enabling a test of H_3. For example, although eCampus and Varsity have nearly the same popularity and page views per user, the SLR is higher for eCampus than for Varsity. We now show empirically that the relative SLRs explain the discounting behavior of the retailers, consistent with our model predictions.

Analysis of Pricing Strategies

To test the expected price relationships under H_1, we use the firms' SLRs along with the pricing data from Table 2. In regression analyses for the 14 retailers studied, we use a price statistic as the dependent variable and SLR as the independent variable. By definition, the resultant standardized coefficient of the SLR and its significance are identical to the Pearson correlation coefficient and its significance. We report these correlations in Table 4.[13] We observe that

---

[12] These book price comparison sites are the top nine sites that identify the booksellers from which they quote prices, as ranked by Google, BookFinder.com, AddAll.com, BestBookDeal.com, AllBookstores.com, TextbookLand.com, AAABookSearch.com, CampusBooks, Compare Shop Books, and ANYBOOKSFORLESS.com. We did not include other general price comparison sites beyond mySimon, because mySimon represents more than 80% of all price comparison activity conducted on the Internet.

[13] The inverse of page views has no significant pricing correlations, but search share (the numerator of SLR) has significant correlations with some pricing variables. We also analyzed the correlations of the pricing statistics with reach and link popularity after omitting Amazon.com and B&N because of their high values. The correlations remain nonsignificant, except for a positive correlation of .64 ($p < .05$) between reach and maximum discount frequency. The page view measures for Amazon.com and Sam Goody may reflect more than book searches by consumers. Delet-

TABLE 4
Correlations of Exploratory Study Variables and Retailer Prices

	Observed Price Promotion Statistics					
Predictor Variables	**Average**	**SD**	**Average Frequency**	**Maximum Frequency**	**Average Depth**	**Maximum Depth**
Reach/millions per day	.051	.093	–.098	.111	–.093	.067
Link popularity	.041	.054	–.097	.107	–.094	.055
Book-link popularity	.042	.051	–.098	.108	–.095	.055
Page views per user	–.040	.258	–.161	.155	–.134	.185
Search share	–.483*	.610**	.435	.628**	.477*	.810***
SLR	–.673***	.503*	.720***	.568**	.744***	.791***
Hypothesis	H_{1a}	H_{1b}	H_{1c}	H_{1d}	H_{1e}	H_{1f}
Predicted relationship to SLR	–	+	+	+	+	+

*$p < .10$.
**$p < .05$.
***$p < .01$.
Notes: N = 14.

the SLR is significantly correlated with the firms' price discounting. The results support H_1, demonstrating that, on average, retailers with a higher SLR have lower prices (H_{1a}) and a higher standard deviation (H_{1b}). They also discount more often (H_{1c} and H_{1d}) and with greater depth (H_{1e} and H_{1f}). We conclude that a firm's SLR is a significant predictor of its promotional activity.

Tests of Retailer Price Distributions

For a given pair of retailers, our model predicts that the retailer with a lower SLR should stochastically dominate the other. The test is based on a distinct rank ordering of the 14 retailers (see Table 3 and Figure 4). However, the SLR proxy based on aggregate retailer data may lack a high degree of precision in distinguishing retailers with comparable SLR values. Therefore, we categorize retailers into three groups with high, medium, or low SLR values and compare retailers in different SLR tiers.[14] The three largest retailers have similar SLR values, ranging from 1.00 to 1.17. Using the spread of .17 as a guideline, 1Bookstreet.com and Page One appear to have SLRs that are relatively close to those of the three large firms, such that it is empirically difficult to classify either firm as having a clearly higher or lower SLR. This observation reveals that four Internet booksellers—Double Discount, Textat-Cost, eCampus, and a1Books—are small retailers with high SLRs; here, we expect P_1 to hold. There are also five small retailers—BookVariety.com, Words Worth Books, Varsity, Worthy, and Sam Goody—with lower SLRs; here, we expect P_2 to hold.

Table 3 indicates the tiers of high, medium, and low SLRs. To test H_2, a retailer in a lower tier of SLR should stochastically dominate retailers in a higher SLR tier. We can examine 65 total pairs of retailers in different tiers for stochastic dominance. Of the 40 retailer pairs whose CDFs do not intersect, the SLR values correctly predict first-order stochastic dominance for 32 pairs (80%), in support of H_2 ($\chi^2 = 14.40$, $p < .01$). We use the deficit functions (integral of the CDFs) to test second-order dominance. Of the 49 retailer pairs whose deficit functions do not intersect, the SLR values correctly predict second-order dominance for 41 pairs (84%), in support of H_2 ($\chi^2 = 22.22$, $p < .01$).[15] For retailers exhibiting stochastic dominance over retailers in the other SLR tiers, the SLR values predict the relative price dispersions.

To test H_3, we examine the price dispersion of small retailers with high or low SLRs in more detail. Table 3 shows four high-SLR small retailers, which H_3 predicts will be stochastically dominated by the five low-SLR small retailers. Of the 20 small retailer pairs, 14 (15) have CDFs (deficit functions) that do not intersect, and the SLR values correctly predict first-order (second-order) stochastic dominance for all retailer pairs ($p < .01$).[16] Thus, H_3 is supported; the SLR can explain significantly whether small firms are high-priced niche players or low-priced discounters.

Exploratory Study Summary

The exploratory study presents evidence that supports the hypotheses, indicating that a retailer's SLR explains its pricing behavior consistent with the model's predictions. The SLR is also adept at explaining whether small firms take a high-priced or low-priced approach. For example, we can explain why Worthy prices high and a1Books prices

---

ing these two retailers does not change the nonsignificance of the page view correlations in Table 4, and the SLR correlations remain significant, except for the standard deviation in prices ($p = .12$).

[14] We thank a reviewer for this suggestion.

[15] As a robustness check, we repeat the analysis putting Books-A-Million in the highest SLR tier and Page One in the low-SLR tier, and again, we find support for H_2 ($p < .01$). We also examine all 91 pairs of the 14 retailers (a nontiered analysis) and again find support for H_2 ($p \leq .01$).

[16] As a robustness check, we repeat the analysis, treating Page One as a low-SLR small retailer, and find support for H_3 ($p < .01$). The results also hold if we place Books-A-Million in the high-SLR tier, though it is not clear whether it is a small retailer. We further examine all 36 pairs of the 9 small retailers (a nontiered analysis) and again find support for H_3 ($p \leq .01$).

low, though both firms are small in the same homogeneous goods market. We also test pricing strategies according to the entire price dispersion curve by using tests of stochastic dominance. The empirical literature has no clear consensus on suitable price dispersion constructs (Pan, Ratchford, and Shankar 2004). The stochastic dominance tests correspond to the predicted theoretical relationships and represent a more complete perspective on retailer price dispersion.

As a summary, consider the microcosm of the three retailers, all dealing in college textbooks: TextatCost, eCampus, and Varsity. From the retailer information given in Table 3, Varsity is the smallest retailer of the three in terms of reach and popularity. Although it experiences similar page views, it receives a lower search share and reach than the other two firms. Standard models and predictions would likely suggest that Varsity take the high-discount approach as the small firm, but the opposite is true. Varsity is the premium-priced retailer among the three. This scenario is possible in our model if Varsity has the lowest SLR, which it does. Furthermore, Varsity carries 89% of the books from our sample that Amazon.com carries (compared with only 57% for eCampus and 48% for TextatCost) and charges higher prices than Amazon.com (see Figure 4). Varsity apparently leverages its low SLR into greater assortment while charging high prices.

Despite the supportive results, the exploratory study's SLR proxy is not ideal, given that it is based on aggregate data. Although page views have been associated with loyalty traits in prior studies, the correlation may be weak for purchasing behavior. The search share measure may not correspond well to retailer-specific switcher sizes, because it is based on price comparison sites that offer a limited perspective on actual switching behavior. Therefore, we conduct a second study using a large panel database of Internet browsing and purchase. This allows for a more microlevel behavioral construction of loyalty and switcher measures, albeit for a more limited number of retailers.

Internet Panel Study

The second study uses the comScore 2002 panel data set to develop switcher and loyal measures. Households in the Internet panel have their Web-browsing and purchase activity recorded. Several recent studies have used comScore data (Danaher 2007; Johnson, Bellman, and Lohse 2003; Moe and Fader 2004a, b; Park and Fader 2004). The data set we use contains 100,000 panelists in the United States from June 2002 to December 2002. Data include site domains visited, number of page views, duration visiting the site, and transaction data if a purchase is made. The data set allows us to formulate switcher and loyalty proxies based on microlevel browsing and purchasing behaviors.

SLR Proxy

We use behavior-based measures to form the SLR proxies. We focus on Internet booksellers in the data set that received purchases; a retailer with only browsing activity would correspond to an undefined SLR. Six retailers for which we have pricing data had purchases in the panel data: 1Bookstreet.com, Amazon.com, B&N, Books-A-Million, eCampus, and Varsity. Book purchases were made by 5304 unique panelists at these retailers over the six-month period.

We take a share-of-purchase perspective (e.g., Tellis 1988) that recognizes Internet retailer loyalty as a lack of search across retailers when making a purchase (Srinivasan, Anderson, and Ponnavolu 2002), consistent with our model's premise.[17] A given purchase occasion corresponds to loyal behavior if no other book retailers or price comparison shopbots are searched before purchase or to switcher behavior if at least one other book retailer or shopbot is searched before purchase.[18] This results in a retailer's SLR measure being the ratio of switcher purchase occasions to loyal purchase occasions. Consistent with prior studies (e.g., Danaher 2007; Moe and Fader 2004a; Park and Fader 2004), we take a daily perspective on panelists' purchase activity in forming the measures.

The retailer data with SLR measures appear in Table 5. In general, there is a high level of loyal purchases. Johnson and colleagues (2004) report that Internet book shoppers in their data set visit approximately 1.2 retailers (compared with 1.25 in our data) and that 70% are loyal to a single site (compared with 79% in our study). The SLR values correlate well with those of the exploratory study for the six retailers (r = .85, p = .03). Note that four of the six retailers correspond to the closely spaced medium level of the

---

[17] A repeat-purchase measure for loyalty is not feasible given the few observations for two of the six retailers.

[18] The price comparison sites checked are those of the search share measure in the exploratory study.

TABLE 5
Retailer Panel Data

Retailer	Total Purchase Visits	Average Number of Retailers Visited	Switcher Purchases	Loyal Purchases	Panel Data SLR
eCampus	72	1.72	37	35	1.06
Books-A-Million	142	1.70	60	82	.73
B&N	1794	1.40	583	1211	.48
Varsity	26	1.23	6	20	.30
1Bookstreet.com	10	1.50	2	8	.25
Amazon.com	5519	1.19	891	4628	.19

Source: comScore (2002; see http://wrds.wharton.upenn.edu).

exploratory SLR (Table 3), and the panel data measure creates a more distinctive spread in the SLR values (Table 5).

Analysis

Table 6 presents correlations of the comScore-based SLR proxies with pricing statistics for the six retailers. The comScore proxy correlates well with the pricing statistics in the expected direction. Even with the small number of retailers, all the correlations are significant, except for the standard deviation in prices. The results support H_1.

Examining the price dispersions among the six retailers, we find that the comScore SLR values correctly predict first- and second-order stochastic dominance for 10 of 13 pairs of nonintersecting price distributions (the distribution of Amazon.com intersects with eCampus and Books-A-Million). Second-order dominance is correctly predicted for 11 of 15 retailer pairs (no intersecting deficit functions). These results support H_2 (first-order $\chi^2 = 3.77$, $p = .05$; second-order $\chi^2 = 3.27$, $p = .07$). Only Amazon.com is out of place in assessing stochastic dominance according to the comScore SLR values. Thus, for the small retailers, the SLR values correctly predict whether they are relatively high or low priced in all cases, consistent with H_3.

Summary of Panel Data Study

The comScore data set allows for a microlevel formulation for a retailer's SLR. Six retailers represent a small sample for analysis, but the results support the model's predictions. An analysis using a retailer's loyal segment size instead of SLR reveals that loyalty alone is incapable of capturing the retailer's pricing strategies. The number of loyals does not correlate significantly with any pricing variable (this result does not change if Amazon.com and B&N are eliminated as large outliers). Loyalty alone also poorly explains the price dispersions, with only 3 of 13 (4 of 15) pairs of retailer distributions correctly predicted for first-order (second-order) stochastic dominance. The results reiterate the important distinction between our segmented-switcher model and those based on loyalty sizes. We also note that analyses using only the first two months of the comScore data set to generate SLR values yield similar results, but 1Bookstreet.com is eliminated because of few switchers or loyals.

The reasonable correlation between the SLR proxies of both studies and the supportive results suggest that a larger retailer sample would generate similar results. Six retailers do not allow for definitive conclusions, but we note that page views from the exploratory study correlate well with the number of loyals from the comScore data (r = .76, $p < .10$), and the search share measure positively correlates with the number of switchers but not significantly (r = .54, n.s.). In total, the studies provide evidence that the SLR can explain pricing strategies, consistent with our model.

Discussion

Price competition among retailers, large and small, involves fundamental decisions about whether and how often to discount prices. Most theoretical models based on asymmetric duopolies suggest that small retailers discount to capture the lucrative switcher market. However, observed price promotion strategies reflect a wide variety of approaches, with some small retailers pricing high and others discounting deeply. Our model examines competitive promotional strategies in an asymmetric oligopoly with segmented switchers. Reflecting the realistic assumption that price-sensitive switchers do not necessarily compare prices at all retailers, switcher segments add an important strategic dimension beyond asymmetric retailer loyalty. Although firms may be classified as large, medium, or small with respect to their loyal segment sizes, this ranking may have little to do with the price-discounting behavior of the firm. We find that the firm's ratio of the relative size of switchers (consumers who consider the retailer's prices) to the relative size of its loyal customers is a better indicator of the firm's price-discounting strategy than loyalty alone. Our model is the first to consider multiple switcher segments, enabling us to categorize a small or large firm as a heavy or light promoter under different loyal and switcher segment compositions. We provide empirical support for our model's key predictions in two studies of Internet booksellers. The segmented-switcher perspective identifies whether a small firm will discount, as is usually predicted, or price high and play the niche. Overall, our model captures a wide variety of price promotion activity, consistent with empirically observed price dispersions. We can summarize our key findings as follows:

- Asymmetry, in terms of different loyal segment sizes and the existence of switcher segments, leads to a multiplicity of discounting strategies. It is possible to categorize when firms will be frequent or deep discounters using the retailers' relative SLRs.

TABLE 6
Correlations of Panel Data SLR Values and Retailer Prices

	Observed Price Promotion Statistics					
	Average	SD	Average Frequency	Maximum Frequency	Average Depth	Maximum Depth
SLR	–.779*	.460	.810*	.839**	.805*	.833**
Hypothesis	H_{1a}	H_{1b}	H_{1c}	H_{1d}	H_{1e}	H_{1f}
Predicted relationship to SLR	–	+	+	+	+	+

*$p < .10$.
**$p < .05$.
Source: comScore (2002; see http://wrds.wharton.upenn.edu).

- A large firm may find it more profitable to offer deeper (shallower) discounts than a smaller firm if the relative share of switchers for which it competes is large (small).
- For midsize firms, a partitioned discounting strategy—one that combines frequent shallow and infrequent deep (or infrequent shallow and frequent deep) discounts—is possible as a result of competition on multiple fronts.
- A small firm may find it profitable to price high and play the niche if the other firms are already intensely competing for the switcher segments the small firm could serve. A small firm may also benefit from strategically choosing to limit its exposure to switchers in the market to reduce the market's overall competitive price intensity.

Using data from the online book market, we empirically validate our model's key implications. Empirical SLR measures from two studies significantly explain price dispersion among retailers, consistent with our model's predictions. The first study uses descriptive retailer statistics to proxy SLR values, and the second study uses a rich set of panel data to determine the number of retailer customers exhibiting loyal or switching behavior. Both studies support the pricing hypotheses that result from our model. Our tests of stochastic dominance are related to the entire price dispersion curve, which is more comprehensive than testing summary price statistics alone.

Managerial Implications

Our results have numerous implications for retailer pricing strategies. Larger firms often face a dilemma due to having both a large loyal following and access to a large number of switchers. The frequency and depth of discounts depend critically on the mix of switcher segments in the overall competitive context. When a vast majority of switchers widely compare prices, the large firm should offer few shallow discounts because other firms will more aggressively discount. In the case in which many small retailers are ignored by the bulk of the switchers, a large firm may find itself more deeply discounting to compete with it nearest rival (e.g., Amazon.com and B&N exhibit moderate levels of discounting). A midsized firm has considerable flexibility in its discounting strategy, such as offering frequent and shallow discounts to compete with the larger retailer but infrequent and deep discounts to compete with smaller firms. The midsized retailer should discount often enough to convince its larger rival to concede some of the switchers, who then behave like loyals to the midsize firm.

Small retailers (even of homogeneous goods) with relatively few loyal customers should not take a deep discount strategy without carefully assessing its access to price-sensitive switchers. If awareness is low or switchers are difficult to reach, a high-priced niche strategy is likely to be more optimal. However, if a wide switcher market is accessible, the depth of discounts depends critically on the specific switcher segments. If most switchers widely compare retailers' prices, in general, the small retailer should offer deep discounts. However, if most switchers concentrate on the larger firms, the small retailer's discounts should be shallower or largely nonexistent if its SLR is small.

The results based on segmented switchers imply certain strategies in how aggressively retailers should reach out to switchers. A small firm that discounts may choose to limit its reach among switchers to some degree to reduce overall price competition. Pursuing fewer switchers can be advantageous for retailers if shallower discounts can be used to capture the switchers that remain. Notably, mySimon considerably reduced the number of firms from which it quoted prices after late June 2002, simultaneously increasing its strength of strategic alliances with its participant stores. By reducing the number of firms returned as the comparison set, mySimon helps these participant stores enjoy higher profits, as our model predicts, while generating higher commissions and referral fees for itself.

The existence of multiple, partially informed switcher segments tends to reduce price competition among the retailers. Retailers must consider carefully not just their loyal customers but also the specific composition of the switcher segments. Acquiring more switchers can be more or less valuable, depending on the segment to which they belong. For example, a midsized retailer would often benefit from targeting its switcher acquisition activities toward its larger rival, given the shallower discounts involved. Overall, the strategies are more than a simple dichotomy of acquisition versus retention because a switcher's value to the retailer depends on its specific switcher segment.

The importance of switcher segmentation suggests that managers should carefully measure switching behaviors. The retailer's value to the switcher and the likelihood of acquiring a switcher customer are both important, but so is the switcher's segment membership. More complete measurement of prior switching behaviors and switcher segment sizes relative to loyals is needed (see Rust, Lemon, and Zeithaml 2004).

Limitations and Further Research

Our model distinguishes between different pricing strategies, such as high-promoter and low-promoter firms, without relying on traffic-building or product differentiation considerations. A segmented-switcher perspective enables us to resolve some varied results found in previous models. We consider this article a unifying extension to stylized game-theoretic models of price comparison and related empirical research. However, we also recognize several limitations. By omitting traffic building from our model, we are unable to observe how segmented switchers alter retailers' traffic-related strategies. Empirically, our proxies for the SLR have several difficulties. Although the exploratory study is supportive, it relies on aggregate descriptive statistics that may not capture microlevel shopping behaviors (see Bucklin and Sismeiro 2003). The study using comScore panel data more directly reflects loyal and switching behaviors based on a retailer's individual customers, but the number of retailers is limited. Future studies could address a variety of product categories to validate our results more robustly. Recent research in trust and loyalty for Internet retailers (e.g., Bart et al. 2005; Shankar, Smith, and Rangaswamy 2003) may suggest other approaches to the loyalty component of SLR.

Endogeneity questions also deserve future study. Retailer pricing strategies may influence the sizes of various switcher and loyal segments as opposed to the exogenous segment sizes in our model. Although our pricing data represent a limited period with no systematic price adjustments, endogenous customer behavior is a potential limitation. Dynamic analyses that capture changes in a person's switcher and loyalty behaviors over time would be fruitful research (for examples on search behavior, see Bucklin and Sismeiro 2003; Johnson, Bellman, and Lohse 2003; Johnson et al. 2004; Moe and Fader 2004a, b).

Further research could also extend recent analyses of Internet shopping agents and firms' strategies regarding the degree of price comparison. The work of Iyer and Pazgal (2003) is an excellent starting point, and our results complement some of their key findings. The existence of segmented switchers and the varied strategies that result give a richer context in which firms can contemplate not only whether to participate in shopping agents but also when to participate in different sites that reach different switcher segments. Our results show that segmented switchers can relax price competition. The implication is that retailers should think not only about how best to invest in loyalty building but also about how they can invest in limiting or expanding switcher segments in relation to other firms in their market. We hope that our theoretical and empirical results motivate further research that more precisely captures not only the retailers' multiple switcher segments but also the pricing strategies that result.

Appendix

Under the model assumptions, the profit functions of the firms are given by the following equation:

$$\pi_i(p_i, p_{-i}) = p_i(n_i + \delta_{12}s_{12} + \delta_{123}s_{123}),$$

where p_i represents the price quoted by Firm i, p_{-i} represents the vector of prices quoted by other firms, and δ_{ij} takes values from the set [0, 1] based on the lowest quoted price. The term δ_{12} applies to Firms 1 and 2 and equals 1 for the lower-priced firm, ½ in the event of a tie in prices, and 0 for the higher-priced firm. The term δ_{123} applies similarly to all three firms, but if two firms quote the lowest price, δ_{123} equals ½ for those firms and 0 for the other firm. For a three-way tie in prices, δ_{123} equals ⅓ for all three firms. We begin with three lemmas that establish properties of the equilibrium solution (proofs are available on request):

Lemma 1: There is no Nash equilibrium in pure strategies.

Lemma 2: Let S_i^*, i = {1, 2, 3}, be the best-response strategy sets (i.e., the set of prices) in the mixed-strategy equilibrium. There is no gap (p_1, p_2) within the joint support $S_1^* \cup S_2^* \cup S_3^*$, where $f_i(p) \equiv 0$ for two or more firms. That is, at least two firms possess positive support at any point within the joint support of prices.

Lemma 3: The CDFs of firms' prices F_i, i = {1, 2, 3}, are continuous, except possibly at r. That is, there are no mass points in the interior or at the lower boundary of the joint support $S_1^* \cup S_2^* \cup S_3^*$.

Proposition 1

P_1 states that when $\phi_3 > \phi_1$, Firms 1 and 3 have mutually exclusive price ranges that, when combined, form Firm 2's price range. Firm 2 competes with Firm 3 in the lower part of its price range and with Firm 1 in the upper part of its price range. The equilibrium CDFs, $F_i(p)$, of the firms' prices are given by the following:

$$F_1(p) = \begin{cases} 0 & p < \underline{p}_1 \\ (p - \underline{p}_1)(n_2 + s_{12})/(ps_{12}) & \underline{p}_1 \le p < r \\ m & p = r \\ 1 & p > r \end{cases},$$

$$F_2(p) = \begin{cases} 0 & p < \underline{p}_2 \\ (p - \underline{p}_2)(n_3 + s_{123})/(ps_{123}) & \underline{p}_2 \le p \le \underline{p}_1 \\ 1 - n_1(r - p)/(ps_{12}) & \underline{p}_1 \le p \le r \\ 1 & p \ge r \end{cases}, \text{ and}$$

$$F_3(p) = \begin{cases} 0 & p < \underline{p}_2 \\ 1 - (\underline{p}_1 - p)(n_2 + s_{12})/(ps_{123}) & \underline{p}_2 \le p \le \underline{p}_1, \\ 1 & p \ge \underline{p}_1 \end{cases}$$

where m is Firm 1's mass point at the reservation price, $\underline{p}_1$ is the lower bound of Firm 1's support, and $\underline{p}_2$ is the common lower bound of Firm 2's and Firm 3's supports given by the following:

$$\underline{p}_1 = \frac{n_1 r(n_2 + s_{12} + s_{123})}{s_{12}(n_2 - n_3 + s_{12}) + n_1(n_2 + s_{12} + s_{123})} > p_1^{min}, \text{ and}$$

$$\underline{p}_2 = \frac{n_1 r(n_2 + s_{12})}{s_{12}(n_2 - n_3 + s_{12}) + n_1(n_2 + s_{12} + s_{123})} \ge p_2^{min}, p_3^{min}.$$

Proof of Proposition 1

From Lemma 1, there is no pure strategy. A detailed exposition of mixed-strategy solution mechanics for models similar to ours is found in the work of Narasimhan (1988). We first define the upper and lower boundaries of the firms' supports. The upper bound of the feasible price set is r. Prices higher than the reservation price will result in no sales, whereas positive profits are possible when the reservation price is quoted. Therefore, the highest price that any firm can charge is the common reservation price r. To determine the lower boundaries, there will likely be two price regions. The lower price region is where only Firms 3 and 2 compete, with an upper bound determined by the lowest price that Firm 1 will quote. The higher price region is determined from this price upward. The minimum price for any firm is when it is indifferent between selling only to its loyals and offering a deep discount to capture the switchers. The minimum prices are $p_i^{min} = n_i r/(n_i + s_{12} + s_{123})$ for Firms 1 and 2 and $p_3^{min} = n_3 r/(n_3 + s_{123})$ for Firm 3. Because $n_1 > n_2 > n_3$, it is always true that $p_2^{min} < p_1^{min}$. Under P_1's assumption that $p_3^{min} < p_1^{min}$, we must consider two cases of whether Firm 3's minimum price is less than or greater than that of Firm 2.

Under the condition that $p_3^{min} < p_2^{min} < p_1^{min}$, Firm 2 can compete with Firm 3 at prices below p_1^{min}, such that both Firms 2 and 3 have positive support below p_1^{min}. Therefore, Firm 1 cannot capture the switcher segments with certainty, even if it prices at p_1^{min}. This conclusion means that the lowest price to which Firm 1 is willing to discount, $\underline{p}_1$, must be higher than $p_1^{\min}$ to balance the prospect of receiving fewer switchers in expectation. Denoting $F_i(p)$ as the cumulative distribution probability of Firm i's prices, we equate Firm 1's profit selling only to its loyals and its profit when also selling to the switchers at $\underline{p}_1$, given the probability that Firms 2 and 3 quote higher prices:

$$n_1 r = n_1 \underline{p}_1 + [1 - F_2(\underline{p}_1)]\underline{p}_1 s_{12} + [1 - F_2(\underline{p}_1)][1 - F_3(\underline{p}_1)]\underline{p}_1 s_{123}$$

$$\Rightarrow \underline{p}_1 = \frac{n_1 r}{n_1 + [1 - F_2(\underline{p}_1)]s_{12} + [1 - F_2(\underline{p}_1)][1 - F_3(\underline{p}_1)]s_{123}}.$$

At prices lower than $\underline{p}_1$, Firms 2 and 3 compete for segment s_{123}, and Firm 2 receives all of segment s_{12} with certainty. Firm 3 has the lowest minimum price, so it will find it profitable to include as a lower support bound only the lowest price Firm 2 will quote. However, because Firm 2 also competes with Firm 1 for s_{12} and Firm 1 has a lower support above $p_1^{\min}$, Firm 2 may find it profitable to adapt a lower support that is higher than $p_2^{\min}$. Thus, Firms 2 and 3 will compete with prices at least up to $\underline{p}_1$. In doing so, they share a common lower bound $\underline{p}_2$, which may be higher than $p_2^{\min}$.

To establish the equilibrium profits, all firms can guarantee the profit $n_i r$ by choosing to price at r. However, in terms of undercutting other firms and serving the switcher segment, Firms 2 and 3 are at an advantage. Firm 3 can improve its profit above $n_3 r$ by pricing at the lowest price any other firm ever discounts, $\underline{p}_2$, and serving s_{123}, which results in a profit of $(n_3 + s_{123})\underline{p}_2$ for Firm 3. Pricing below $\underline{p}_2$ is never optimal for Firm 3, because it could then raise its price and still capture s_{123} for sure. Similarly, Firm 2 can improve its profit above $n_2 r$ by pricing at the minimum price Firm 1 will ever feasibly reduce its price to, $\underline{p}_1$, and serving s_{12}, which results in a profit of $(n_2 + s_{12})\underline{p}_1$ for Firm 2. The equilibrium solution reveals that Firm 2 would never deviate and price below the lower bound $\underline{p}_2$, because it would lose more profit from selling at a lower price to s_{12} and n_2 than it could gain by capturing s_{123} with higher probability.

When Firms 2 and 3 compete with prices between $\underline{p}_2$ and $\underline{p}_1$, both have lower prices than Firm 1 with a probability of 1. Thus, either one can capture the switcher segment s_{123} if it can price lower than the other. The competition in this interval is similar to Narasimhan's (1988) base model, with two exceptions. First, the upper limit of the interval is not r, but rather $\underline{p}_1$, which is less than r. Second, Firm 2 also considers that whenever it prices in $[\underline{p}_2, \underline{p}_1]$, it will serve the switcher segment s_{12} with a probability of 1. Firms 2 and 3 will randomize their prices in this interval so that the expected profit will be equal to their equilibrium profits. We can write the equilibrium conditions for the interval $[\underline{p}_2, \underline{p}_1]$, with the exception of $\underline{p}_1$, as follows:

(A1) $E\pi_2 = (n_2 + s_{12})\underline{p}_1 = n_2 p + s_{12} p + [1 - F_3(p)]ps_{123}$, and

(A2) $E\pi_3 = (n_3 + s_{123})\underline{p}_2 = n_3 p + [1 - F_2(p)]ps_{123}$

Note that in Equation A1, Firm 2 serves its loyal segment and switcher segment s_{12} with any price it quotes. It serves switcher segment s_{123} only if its price is lower than Firm 3's price. In Equation A2, Firm 3 serves its loyal segment with any price it quotes, but it serves switcher segment s_{123} only if its price is lower than Firm 2's price. Although attempts to price low in this interval increase Firm 2's chances of serving s_{123}, they also decrease Firm 2's guaranteed profit from s_{12}. Thus, because $p_3^{\min} < p_2^{\min} < \underline{p}_2$, both firms will share $\underline{p}_2$ as the lower bounds of their supports, as in the work of Narasimhan (1988). This also means that neither firm will have a mass point at $\underline{p}_2$, because in that case, the other firm would have a motivation to undercut $\underline{p}_2$. Therefore, $F_2(\underline{p}_2)$ and $F_3(\underline{p}_2)$ both equal zero.

To solve Equations A1 and A2, we also need to specify values for $\underline{p}_2$ and $\underline{p}_1$. Note that $\underline{p}_1$ is the lowest price Firm 1 will quote. Thus, the profit Firm 1 makes with the price $\underline{p}_1$ must be equal to its expected profit:

(A3) $$E\pi_1 = n_1 r = n_1 \underline{p}_1 + [1 - F_2(\underline{p}_1)]\underline{p}_1 s_{12} + [1 - F_2(\underline{p}_1)][1 - F_3(\underline{p}_1)]\underline{p}_1 s_{123}.$$

By solving Equations A1, A2, and A3 simultaneously, in combination with the cumulative distribution conditions at $\underline{p}_2$, we find the solutions to this set of equilibrium conditions in the interval $[\underline{p}_2, \underline{p}_1]$, with the exception of $\underline{p}_1$, as follows:

(A4) $$F_2(p) = \frac{(p - \underline{p}_2)(n_3 + s_{123})}{ps_{123}},$$

(A5) $$F_3(p) = 1 - \frac{(\underline{p}_1 - p)(n_2 + s_{12})}{ps_{123}},$$

(A6) $$\underline{p}_1 = \frac{n_1 r(n_2 + s_{12} + s_{123})}{s_{12}(n_2 - n_3 + s_{12}) + n_1(n_2 + s_{12} + s_{123})}, \text{ and}$$

(A7) $$\underline{p}_2 = \frac{n_1 r(n_2 + s_{12})}{s_{12}(n_2 - n_3 + s_{12}) + n_1(n_2 + s_{12} + s_{123})}.$$

On the basis of this solution set, we also observe that $F_3(\underline{p}_1)$ equals 1, thus showing that in the next interval, upward of $\underline{p}_1$, only Firms 1 and 2 will compete. Furthermore, the solution holds as long as $\underline{p}_2 < p_1^{\min}$, because otherwise, Firm 1 would have an incentive to discount below $\underline{p}_1$. For $n_1 \geq n_2 + s_{12} + n_3(s_{12}/s_{123})$, $\underline{p}_2$ will never exceed $p_1^{\min}$. In general, Lemma 2 proves that there are no gaps in the support. Lemma 3 proves continuity in the distribution functions.

To solve for the pricing behavior in the next support interval $[\underline{p}_1, r]$, we write the equilibrium conditions, with the exception of r, as follows:

(A8) $E\pi_1 = n_1 r = n_1 p + [1 - F_2(\underline{p}_1)][1 - \bar{F}_2(p)]ps_{12}$, and

(A9) $E\pi_2 = (n_2 + s_{12})\underline{p}_1 = n_2 p + [1 - F_1(p)]ps_{12}$,

where $\bar{F}_2(p)$ is the conditional CDF given that $p > \underline{p}_1$ for Firm 2. Firm 2's unconditional CDF in the interval $[\underline{p}_1, r]$ is given by $F_2(\underline{p}_1) + [1 - F_2(\underline{p}_1)]\bar{F}_2(p)$. The solution to this set of equations gives the final results, where the mass point for Firm 1 at r is a straightforward calculation, given $F_1(p)$:

$$m = \frac{n_1(n_2 + s_{12} + s_{123}) - n_2(n_2 - n_3 + s_{12})}{s_{12}(n_2 - n_3 + s_{12}) + n_1(n_2 + s_{12} + s_{123})}.$$

This concludes the solution for $p_3^{min} < p_2^{min} < p_1^{min}$.

For the second case of $p_2^{min} \leq p_3^{min} < p_1^{min}$ (i.e., $\phi_2 \geq \phi_3 > \phi_1$), the mixed-strategy equilibrium is identical to that described previously for $p_3^{min} < p_2^{min} < p_1^{min}$ (the proof is available on request). Q.E.D.

Proposition 2

P_2 states that when $\phi_3 \leq \phi_1$, only Firms 1 and 2 discount, whereas Firm 3's price is set at r. The resultant CDFs are as follows:

$$F_1(p) = \begin{cases} 0 & p < p_1^{min} \\ 1 - \dfrac{p_1^{min}(n_2 + s_{12} + s_{123}) - n_2 p}{p(s_{12} + s_{123})} & p_1^{min} \leq p < r \\ \dfrac{n_1 - n_2 \cdot}{n_1 + s_{12} + s_{123}} & p = r \\ 1 & p > r \end{cases},$$

$$F_2(p) = \begin{cases} 0 & p < p_1^{min} \\ 1 - \dfrac{n_1(r - p)}{p(s_{12} + s_{123})} & p_1^{min} \leq p \leq r, \text{ and} \\ 1 & p \geq r \end{cases}$$

$$F_3(p) = \begin{cases} 0 & p < r \\ 1 & p \geq r \end{cases}.$$

Proof of Proposition 2

Under $p_2^{min} < p_1^{min} \leq p_3^{min}$, Firm 3 has the highest minimum price and thus has no advantage in offering deep enough discounts to serve the switcher segment s_{123}. However, Firms 2 and 3 will be competing for the business of s_{12} and s_{123}. Compared with Firm 1's position in P_1, Firm 3 is essentially a high-priced niche player that serves only its loyals. Firm 3 will not serve segment s_{123}, because both Firm 1 and Firm 2 quote lower prices with positive probability as they compete for s_{12}. Thus, the candidate equilibrium profits for Firms 1, 2, and 3 are $n_1 r$, $(n_2 + s_{12} + s_{123})p_1^{min}$, and $n_3 r$, respectively. Although Firm 2 can offer the lowest price, it will never discount below p_1^{min}, the price at which it can successfully capture both switcher segments. In this case, Firms 1 and 3 have equilibrium profits they would make from sales to their loyal segments only. For Firms 1 and 2, we solve as follows:

(A10) $$E\pi_1 = n_1 r = n_1 p + [1 - F_2(p)]p(s_{12} + s_{123}), \text{ and}$$

(A11) $$E\pi_2 = (n_2 + s_{12} + s_{123})p_1^{min} = n_2 p + [1 - F_1(p)]p(s_{12} + s_{123}).$$

The solution to this set is as follows:

(A12) $$F_1(p) = 1 - \frac{p_1^{min}(n_2 + s_{12} + s_{123}) - n_2 p}{p(s_{12} + s_{123})}, \text{ and}$$

(A13) $$F_2(p) = 1 - \frac{n_1(r - p)}{p(s_{12} + s_{123})}.$$

To show that Firm 3 will indeed never discount in equilibrium, the lowest price Firm 3 will ever quote is p_3^{min}. At this price, its expected profit is represented by the following:

(A14) $$\pi_3(p_3^{min}) = n_3 p_3^{min} + [1 - F_2(p_3^{min})][1 - F_1(p_3^{min})]p_3^{min} s_{123}.$$

Inserting values from Equations A12 and A13 into Equation A14, we indeed find that $\pi_3(p_3^{min}) < n_3 r$, which means that Firm 3 will never price at p_3^{min}, given that Firms 1 and 2 are already competing for the switcher segments below this price. In addition, because it is only Firms 1 and 2 that can compete at p_3^{min}—and possibly above, as we have just shown—the CDFs presented by Equations A12 and A13 will also remain valid above p_3^{min}. We can solve for the lowest price point above p_3^{min} to which Firm 3 will ever reduce its price by solving the following:

(A15) $$n_3 r = n_3 p + [1 - F_1(p)][1 - F_2(p)]p s_{123}.$$

The only solution to this equation that rises above p_3^{min} is r. Thus, given that Firms 1 and 2 are already competing for the switcher segment below p_3^{min}, Firm 3 will never price in the interval $[p_3^{min}, r]$. Rather, it will only price at r. Thus, the solutions to Equations A12 and A13 remain valid until r. With this solution, Firm 1 also has a mass point at r that equals $(n_1 - n_2)/(n_1 + s_{12} + s_{123})$. Although both Firms 1 and 3 have positive masses at r, because Firm 2 has a lower price than r with a probability of 1, Firms 1 and 3 will receive their guaranteed profits $n_i r$, whereas Firm 2's profit will be higher than $n_2 r$. Q.E.D.

REFERENCES

Allen, Gove and Jianan Wu (2002), "Data Collection from Intelligent Agents: When Is Enough, Enough?" working paper, Freeman School of Business, Tulane University.

Bart, Yakov, Venkatesh Shankar, Fareena Sultan, and Glen Urban (2005), "Are the Drivers and Role of Online Trust the Same for All Web Sites and Consumers? A Large-Scale Exploratory Empirical Study," *Journal of Marketing*, 69 (October), 133–52.

Baye, Michael R., Dan Kovenock, and Casper G. de Vries (1992), "It Takes Two to Tango: Equilibria in a Model of Sales," *Games and Economic Behavior*, 4 (4), 493–510.

——— and John Morgan (2001), "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets," *American Economic Review*, 91 (3), 454–74.

Blattberg, Robert and John Deighton (1991), "Interactive Marketing: Exploiting the Age of Addressability," *Sloan Management Review*, 33 (1), 5–14.

Bolton, Ruth N. and Venkatesh Shankar (2003), "An Empirically Derived Taxonomy of Retailer Pricing and Promotion Strategies," *Journal of Retailing*, 79 (4), 213–24.

Bucklin, Randolph E. and Catarina Sismeiro (2003), "A Model of Web Site Browsing Behavior Estimated on Clickstream Data," *Journal of Marketing Research*, 40 (August), 249–67.

Burdett, Kenneth and Kenneth L. Judd (1983), "Equilibrium Price Dispersion," *Econometrica*, 51 (4), 955–69.

Carlton, Dennis W. and Judith A. Chevalier (2001), "Free Riding and Sales Strategies for the Internet," *Journal of Industrial Economics*, 49 (4), 441–61.

Chen, Yuxin, Chakravarthi Narasimhan, and Z. John Zhang (2001), "Individual Marketing with Imperfect Targetability," *Marketing Science*, 20 (1), 23–41.

Clay, Karen, Ramayya Krishnan, and Eric Wolff (2001), "Prices and Price Dispersion on the Web: Evidence from the Online Book Industry," *Journal of Industrial Economics*, 49 (4), 521–39.

Danaher, Peter J. (2007), "Modeling Page Views Across Multiple Websites with an Application to Internet Reach and Frequency Prediction," *Marketing Science*, 26 (3), 422–37.

———, Guy W. Mullarkey, and Skander Essegaier (2006), "Factors Affecting Web Site Visit Duration: A Cross-Domain Analysis," *Journal of Marketing Research*, 43 (May), 182–94.

Demers, Elizabeth and Baruch Lev (2001), "A Rude Awakening: Internet Shakeout in 2000," *Review of Accounting Studies*, 6 (2–3), 331–69.

Fox, Edward J. and Stephen J. Hoch (2005), "Cherry-Picking," *Journal of Marketing*, 69 (January), 46–62.

Gibbons, Robert (1992), *Game Theory for Applied Economists*. Princeton, NJ: Princeton University Press.

Goldfarb, Avi (2002), "Online Marketing Metrics and Financial Performance," working paper, Department of Economics, Northwestern University.

Hoch, Stephen J., Xavier Drèze, and Mary E. Purk (1994), "EDLP, Hi-Lo, and Margin Arithmetic," *Journal of Marketing*, 58 (October), 16–27.

Iyer, Ganesh and Amit Pazgal (2003), "Internet Shopping Agents: Virtual Co-Location and Competition," *Marketing Science*, 22 (1), 85–106.

Johnson, Eric J., Steven Bellman, and Gerald L. Lohse (2003), "Cognitive Lock-In and the Power Law of Practice," *Journal of Marketing*, 67 (April), 62–75.

——— Wendy W. Moe, Peter S. Fader, Steven Bellman, and Gerald L. Lohse (2004), "On the Depth and Dynamics of Online Search Behavior," *Management Science*, 50 (3), 299–308.

Koçaş, Cenk (2002), "Evolution of Prices in Electronic Markets Under Diffusion of Price-Comparison Shopping," *Journal of Management Information Systems*, 19 (3), 99–119.

——— and Tunga Kiyak (2006), "Theory and Evidence on Pricing by Asymmetric Oligopolies," *International Journal of Industrial Organization*, 24 (1), 83–105.

Lal, Rajiv and Miklos Sarvary (1999), "When and How Is the Internet Likely to Decrease Price Competition?" *Marketing Science*, 18 (4), 485–503.

Leshno, Moshe and Haim Levy (2002), "Preferred by 'All' and Preferred by 'Most' Decision Makers: Almost Stochastic Dominance," *Management Science*, 48 (8), 1074–1085.

Moe, Wendy W. and Peter S. Fader (2004a), "Capturing Evolving Visit Behavior in Clickstream Data," *Journal of Interactive Marketing*, 18 (1), 5–19.

——— and ——— (2004b), "Dynamic Conversion Behavior at E-Commerce Sites," *Management Science*, 50 (3), 326–35.

Montgomery, Alan L., Karthik Hosanagar, Ramayya Krishnan, and Karen D. Clay (2004), "Designing a Better Shopbot," *Management Science*, 50 (2), 189–202.

Narasimhan, Chakravarthi (1988), "Competitive Promotional Strategies," *Journal of Business*, 61 (4), 427–49.

Pan, Xing, Brian T. Ratchford, and Venkatesh Shankar (2004), "Price Dispersion on the Internet: A Review and Directions for Future Research," *Journal of Interactive Marketing*, 18 (4), 116–35.

Park, Young-Hoon and Peter S. Fader (2004), "Modeling Browsing Behavior at Multiple Websites," *Marketing Science*, 23 (3), 280–303.

Rajiv, Surendra, Shantanu Dutta, and Sanjay K. Dhar (2002), "Asymmetric Store Positioning and Promotional Advertising Strategies: Theory and Evidence," *Marketing Science*, 21 (1), 74–96.

Raju, Jagmohan, V. Srinivasan, and Rajiv Lal (1990), "The Effects of Brand Loyalty on Competitive Price Promotional Strategies," *Management Science*, 36 (3), 276–304.

Rust, Roland T., Katherine N. Lemon, and Valarie A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," *Journal of Marketing*, 68 (January), 109–127.

Salop, Steven and Joseph Stiglitz (1977), "Bargains and Rip-Offs: A Model of Monopolistically Competitive Price Dispersion," *Review of Economic Studies*, 44 (3), 493–510.

Shankar, Venkatesh and Ruth N. Bolton (2004), "An Empirical Analysis of Determinants of Retailer Pricing Strategy," *Marketing Science*, 23 (1), 28–49.

———, Amy K. Smith, and Arvind Rangaswamy (2003), "Customer Satisfaction and Loyalty in Online and Offline Environments," *International Journal of Research in Marketing*, 20 (2), 153–75.

Smith, Michael and Erik Brynjolfsson (2001), "Consumer Decision-Making at an Internet Shopbot: Brand Still Matters," *Journal of Industrial Economics*, 49 (4), 541–58.

Srinivasan, Srini S., Rolph Anderson, and Kishore Ponnavolu (2002), "Customer Loyalty in E-Commerce: An Exploration of Its Antecedents and Consequences," *Journal of Retailing*, 78 (1), 41–50.

Tellis, Gerard J. (1988), "Advertising Exposure, Loyalty, and Brand Purchase: A Two-State Model of Choice," *Journal of Marketing Research*, 25 (May), 134–44.

Urbany, Joel E., Peter R. Dickson, and Rosemary Kalapurakal (1996), "Price Search in the Retail Grocery Market," *Journal of Marketing*, 60 (April), 91–104.

Varian, Hal R. (1980), "A Model of Sales," *American Economic Review*, 70 (4), 651–59.

JOURNALS OF THE AMERICAN MARKETING ASSOCIATION AVAILABLE IN JSTOR

Since 2003, the American Marketing Association has worked with JSTOR, a not-for-profit online digital archive, to provide access to complete back runs of AMA journals and preserve them in a long-term archive. Users may currently search, browse, view, and print full-text PDF versions of articles in two AMA journals, from their first year published through all but the most recent five volumes:

Journal of Marketing (1936)
Journal of Marketing Research (1964)

These journals are available as part of JSTOR's Arts & Sciences Complement and Business II Collection. Users at institutions that participate in any of these collections can access the publications directly at ‹www.jstor.org›. The AMA publications *Journal of Public Policy & Marketing* and *Journal of International Marketing* will also be available in the JSTOR archive shortly.

Individual articles from the back issues of AMA's publications are available for purchase at $24.00 per article through JSTOR's Publisher Sales Service, a collaborative initiative to facilitate access to important scholarly literature. This program is meant to enable independent scholars unaffiliated with a participating institution or publisher to access single articles from journals archived in JSTOR. For more information on this program, please visit ‹http://www.jstor.org/about/pss_title_list.html›.

As the industry leader in providing expertise and knowledge to marketing professionals, the American Marketing Association is proud to collaborate with JSTOR to ensure the preservation of and access to the important scholarly literature of our field.

JSTOR is a not-for-profit organization helping the scholarly community discover, use and build upon a wide range of intellectual content in a trusted digital archive. Our efforts provide for the long-term preservation and access of leading academic journals and scholarly literature from around the world. Our work is supported by libraries, scholarly societies, publishers, and foundations.

Information regarding JSTOR available at www.jstor.org
149 Fifth Avenue, New York, NY 10010

NEW
MARKETING
TEXTBOOK

Don't make your students choose between:

- New textbook, price $150 or more
- Second hand textbook

*Assign them a **great textbook** at an **affordable** price*

$45 hard copy / **$25** electronic

Managing Marketing in the 21st Century

Examine ***Managing Marketing in the 21st Century*** online at **www.mm21c.com** (instructors' area available)

Noel Capon (author of *The Marketing Mavens*)
with
James Mac Hulbert
Columbia Business School

Wessex

www.mm21c.com